Characterisation of HIV-1 Subtype C Envelope Functional Determinants of Dual Infected Individuals

By

Shatha Sultan Ahmed Omar

Thesis Presented for the Degree of

DOCTOR OF PHILOSOPHY

Department of Integrative Biomedical Sciences

UNIVERSITY OF CAPE TOWN

February 2018

DECLARATION

I, Shatha Sultan Ahmed Omar, hereby declare that the work on which this dissertation/thesis is based is my original work (except where acknowledgements indicate otherwise) and that neither the whole work nor any part of it has been, is being, or is to be submitted for another degree in this or any other university.

I empower the university to reproduce for the purpose of research either the whole or any portion of the contents in any manner whatsoever.

Signature: Signed by candidate

Date: 19/02/2018

DEDICATION

After an intensive period of seven years, I would like to dedicate this work to the soul of my father Sultan Ahmed Omer and my mother Zahra Hibatulla Ali “the minaret of education” that enlighten my path of study from childhood up to higher degree. And to the soul of my kind husband who has been supporting me all over the time and particularly, during my stay in South Africa and encouraged me fulfilling my daily experimental and research duties.

ACKNOWLEDGEMENTS

First and foremost, I have to thank and express my gratefulness to “Allah” the almighty God who provided me with strength throughout my social and scientific life.

I would like to express the deepest gratefulness to my committee chair Dr. Zenda Woodman, who continually conveyed a spirit of exploration in regard to research and scholarship. Without her aspiring guidance, invaluably constructive criticism, friendly advice and persistent help this dissertation would not have been possible.

My special thanks go to Dr. David Micklem, Director of Biomarkers and Assay Development at BerGenBio ASA, for his permanent and tremendous help that introduced me to charity work in this world.

I would also like to thank Dr. Philippe Selhorst and Dr. Melissa Abraham for their kind training and assistance in the replication assay.

I am thankful to the Organisation for Women in Science for the Developing World (OWSD), Schlumberger foundation, Faculty for the Future (FFTF) and Poliomyelitis Research Foundation (PRF) for their financial support that I received from them without which my studies would not have been possible.

I would like to show gratitude to my bright daughter who has been sharing with me the ups and downs and the happy and sad moments of this achievement.

Getting through my dissertation required more than academic support, and I have many, many people to thank for listening to and, at times, having to tolerate me over the past seven years. For you, relatives and friends, I cannot begin to express my gratitude and appreciation being in my life.

TABLE OF CONTENTS

DECLARATION I
DEDICATION II
ACKNOWLEDGEMENTS III
TABLE OF CONTENTS IV
LIST OF TABLES VIII
LIST OF FIGURES IX
ABBREVIATIONS XII
ABSTRACT 1
CHAPTER 1 LITERATURE REVIEW. 4
1.1 Introduction 4
1.2 Overall HIV structure 4
1.3 HIV Diversity 6
1.4 Mechanisms and impact of Diversification 8
1.4.1 Point mutation 9
1.4.2 HIV-1 Recombination 10
1.5 HIV-1 transmission 12
1.6 HIV-1 disease progression 15
1.7 Dual infection 18
1.7.1 Incidence of dual infection 19
1.7.2 Method of Identification 22
1.7.3 Association between dual infection and increased disease progression 23
1.7.4 Dual infection and Drug Resistance 25
1.8 HIV-1 Envelope 26
1.8.1 Structure and function 26
1.9 Viral fitness 36
1.9.1 Env as an HIV fitness determinant 36
1.9.2 Env fitness and disease progression 37
1.10 Research Aim and objectives 38
1.10.1 Rational 38
1.10.2 Objectives 38

CHAPTER 2. THE ROLE OF HIV-1 SUBTYPE C ENVELOPE ENTRY EFFICIENCY IN INVIVO OUTGROWTH OF VARIANTS INFECTED DUAL INFETCED INDIVIDUALS. 40
2.1 Introduction 40
2.2 Research Aim and Objectives 42
2.3 Materials and Methods 43
2.3.1 Study cohort and samples used in this analysis 43
2.3.2 Analysis of SGA-derived envelope sequences 44
2.3.3 Cloning functional Env SGA 45
2.3.4 Pseudovirus Entry Efficiency 47
2.3.5 Env determinants for viral fitness 49
2.3.6 Site direct mutagenesis 50
2.4 Statistical analysis 53
2.5 Results 54
2.5.1 Diversity of variants infecting dual infected individuals 54
2.5.2 Frequency of viral populations infecting dual-infected individuals 56
2.5.3 Identification of phylogenetically distinct master sequences 56
2.5.4 Changes in Env Entry efficiency of viruses infecting dual infected individuals over time 69
2.5.5 Env Fitness determinants 80
2.6 Discussion 87
CHAPTER 3. LONGITUDINAL, PHENOTYPIC CHARACTERISATION OF HIV-1 SUBTYPE C ENVELOPE ISOLATED FROM DUAL INFECTED INDIVIDUALS. 93
3.1 Introduction 93
3.2 Research Aim and Objectives 95
3.3 Materials and Methods 96
3.3.1 Samples used in this study 96
3.3.2 Coreceptor phenotype 96
3.3.3 Cellular tropism 96
3.3.4 Env sensitivity to entry inhibitor T-20 99
3.3.5 Expression and incorporation of Envelope into pseudovirus 99
3.3.6 Data analysis 100
3.3.7 Statistical analysis 101
3.4 Results 103

3.4.1 Coreceptor phenotype 103
3.4.2 Differences in cellular tropism 103
3.4.3 Validity of TZM-bl cells to compare Pseudovirion entry efficiency 107
3.4.4 Pseudovirus Fusion Capacity 112
3.4.5 Association between PSV entry efficiency and fusion capacity 116
3.4.6 Envelope Expression and cleavage 117
3.4.7 Env incorporation 121
3.4.8 Characterisation of Env fitness determinants 126
3.4.9 Summary of results 135
3.5 Discussion 136
CHAPTER 4. UNDERSTANDING THE EFFECT OF ENVELOPE ON VIRAL REPLICATION. 142
4.1 Introduction 142
4.2 Research Aim and Objectives 144
4.3 Material and Methods 145
4.3.1 Samples 145
4.3.2 Plasmids used in the assay 145
4.3.3 Yeast recombination assay 147
4.3.4 Generation of Infectious Molecular Clones 149
4.3.5 Virus Titre 149
4.3.6 IMC Replication 149
4.3.7 Statistical analysis 151
4.4 Results 152
4.4.1 Generation of Infectious Molecular Clones 152
4.4.2 Identification and isolation of responsive donor PBMCs 152
4.4.3 Replication of IMCs in PBMCs 154
4.4.4 Association between viral replication in PBMCs and in vivo frequency of variants 157
4.4.5 Association between viral replication in PBMCs and PSV Entry Efficiency 158
4.4.6 Association between viral replication in PBMCs and Env fusion capacity 162
4.4.7 Association between viral replication in PBMCs and disease progression 164
4.4.8 The impact of Env fitness determinants on IMC replication 167

4.4.9 Summary of results 170
4.5 Discussion 171
CHAPTER 5. SUMMARY AND CONCLUSION. 174
5.1 Discussion 174
5.2 Conclusion 182
5.3 Summary of results 183
BIBILOGRAPHY 185
APPENDIX A. PRIMERS SEQUENCES AND PCR CONDITIONS 216
APPENDIX B. ALIGNMENTS AND PHYLOGENETIC TREES 219
APPENDIX C. MEDIA AND SOLUTIONS 226

LIST OF TABLES

Table 1.1 Incidence of dual infections.....21
Table 2.1. Characterisation of viral populations over time in dual infected individuals.....64
Table 3.1. Density of CD4 and CCR5 induced by Minocycline and Ponesterone A 98
Table 3.2. Antibodies used in Western blot.....101
Table 3.3. Ranking of PSV entry efficiency according to cell type 107
Table 3.4. Overall relationship between PSV entry efficiency, in vivo viral outgrowth and Envelope function and processing.....135
Table 3.5. Phenotypic characterisation of Envelope fitness determinants 135
Table 4.1. Ranking of replication capacity of IMCs in PBMCs.....170
Table 4.2. Summary of associations between infectious molecular clone replication capacity and *in vivo* outgrowth and other Env phenotypes 170
Table 4.3 Impact of Env fitness determinants on chimeric infectious molecular clone replication capacity 170
Table 5.1. Summary of Env phenotype associated with *in vivo* outgrowth overtime 183
Table 5.2. Association between fitness determinants and Env phenotype 184

LIST OF FIGURES

Figure 1.1. HIV-1 genes and proteins. 5
Figure 1.2. HIV Classification. 7
Figure 1.3. Distribution of HIV-1 subtypes worldwide. 8
Figure 1.4. Template switching during HIV-1 reverse transcription. 11
Figure 1.5. Recombination between 2 HIV viruses can select for fitter viruses. 12
Figure 1.6. HIV-1 transmission bottleneck. 14
Figure 1.7. Time course of HIV-1 infection and disease progression. 18
Figure 1.8. Schematic of Envelope Structure. 28
Figure 1.9. HIV entry process. 29
Figure 1.10. HIV-1 Env trimer. 32
Figure 1.11. Location of broadly neutralizing antibody (bNAbs) epitopes on the Envelope trimer. 33
Figure 1.12. Env incorporation models explains. 35
Figure 2.1. CAP84 chimeric Env clones and mutants. 51
Figure 2.2. Construction of CAP267 chimeras and CAP137 mutant. 52
Figure 2.3. Diversity and Diversification of envelope sequence over time. 55
Figure 2.4. Phylogenetic analysis of SGA-derived *env* at early time points. 58
Figure 2.5. Highlighter Plot of sequences generated at early time points. 59
Figure 2.6. Sequence diversity over the course of infection. 62
Figure 2.7. RIP analysis to determine whether AB variants are recombinants. 63
Figure 2.8. In vivo frequency of viral populations over time. 68
Figure 2.9. Pseudovirion Entry Efficiency over 12 months of infection. 75
Figure 2.10. Overall relationship between "*in vivo*" frequency and Env entry efficiency over time. 77
Figure 2.11. Env entry efficiency and disease markers (viral load and CD4 count) over time. 79
Figure 2.12. Association between pseudovirion entry efficiency and CD4+ T cell count. 80

Figure 2.13. Entry efficiency of CAP84 Envelope chimeras and mutants. 84
Figure 2.14. Identification of CAP137 Envelope determinants.....85
Figure 2.15. Entry efficiency of CAP267 chimeras.....86
Figure 3.1. Determine coreceptor tropism over time. 104
Figure 3.2. Cellular tropism using dual inducible 293 Affinofile cells. 106
Figure 3.3. PSV entry efficiency in U87 CD4+CCR5+ cell line.....108
Figure 3.4. PSV Entry Efficiency in Affinofile cells expressing $CD4^{high}/CCR5^{low}$. 109
Figure 3.5. Pseudovirus entry efficiency relative to changes in CCR5 levels. 110
Figure 3.6. Correlation between PSV entry efficiency of TZM-bl and T-lymphocyte like Affinofile cells. 111
Figure 3.7. Pseudovirion sensitivity to the entry inhibitor, T20. 115
Figure 3.8. Correlation between pseudovirus entry efficiency and fusion capacity. 116
Figure 3.9. CAP137 Env Expression. 118
Figure 3.10. CAP267 Env Expression. 120
Figure 3.11. Correlation of gp160 cleavage with Env entry efficiency and fusion capacity. 121
Figure 3.12. CAP137 Env incorporation into PSVs. 123
Figure 3.13. CAP267 Env incorporation into PSVs. 124
Figure 3.14. Correlation between Env gp120 expression and gp120 incorporation.....125
Figure 3.15. Correlation between Env entry efficiency and gp120 incorporation. 126
Figure 3.16. Fusion capacity of CAP84 and CAP267 chimeras. 128
Figure 3.17. Env Expression and cleavage of CAP267 chimeras. 130
Figure 3.18. CAP267 Env incorporation of the chimeras. 131
Figure 3.19. CAP137 Env Expression. 133
Figure 3.20. CAP137 mutant incorporation. 134
Figure 3.21. Structural modelling of potential N-glycans at positions N332 and N339. 141
Figure 4.1. Map of shuttle vector and helper plasmid used in yeast recombination assay. 146
Figure 4.2. Replication in PBMCs from different donors. 153

Figure 4.3. Replication of CAP137 IMCs. 155
Figure 4.4. Replication of CAP267 chimeric IMCs. 156
Figure 4.5. Association between Replication Capacity and in vivo frequency of viral populations over time. 158
Figure 4.6. Association between overall outgrowth of viral populations and Env function. 161
Figure 4.7. Correlation between IMC Replication capacity and PSV entry efficiency.....162
Figure 4.8. Correlation between RC and fusion capacity. 164
Figure 4.9. Association between replication capacity of chimeric infectious molecular clones and markers of disease progression. 166
Figure 4.10. Impact of CAP137 PNG at position 332 on infectious molecular clone replication capacity. 168
Figure 4.11. Replication Capacity of CAP267 chimeras. 169

ABBREVIATIONS

AIDS	Acquired immune deficiency syndrome
ART	Antiretroviral therapy
AUC	Area under the curve
BCN	Broadly cross neutralising antibodies
bnAbs	Broadly neutralising antibodies
C1-C5	Conserved regions
CAPRISA	Centre for the AIDS Programme of Research in South Africa
CCR5	C-C Motif Chemokine Receptor
CMV	Cytomegalovirus
CO2	Carbon dioxide
CRFs	Circulating recombinant forms
CSM	Complete supplement mixture
CT	C- terminal cytoplasmic tail
CTL	Cytotoxic T-cell
CTL	Cytotoxic T-lymphocytes
CXCR4	C-X-C chemokine receptor type 4
DCs	Dendritic cells
DC-SIGN	Dendritic cell–specific intercellular adhesion molecular 3-grabbing non-integrin
DMEM	Dulbecco's minimal essential medium
E.coli	*Escherichia coli*
EC	Elite controllers
ECL2	Extracellular loop-2
EDTA	Ethylenediaminetetraacetic acid
ELISA	Enzyme-linked immunosorbent assay
ENF	Enfuvirtide
Env	Envelope
FACS	Fluorescence-activated cell sorting
FCS	Fetal calf serum
gag	Coding for the viral capsid proteins
GALT	Gut-associated lymphoid tissue
gp120	120kDa Envelope Glycoprotein
gp41	41kDa Envelope Glycoprotein
HAART	Highly active antiretroviral therapy

HEK-293T	Human embryonic kidney cells
HIV-1	Human immunodeficiency virus-type1
HIV-2	Human immunodeficiency virus-type 2
HLA	Human leukocyte antigen
HMA	Heteroduplex mobility assay
HR1 and HR2	Heptad repeat regions1 and 2
HTA	Heteroduplex tracking assay
IL-2	Interleukin 2
IMC	Infectious molecular clone
Ks	Kennedy sequence
LB	Luria broth
LEU	Leucine
LL	Dileucine motif
LTNPs	Long-term non progressors
M	Molar
MCS	Multiple cloning site
MDR	Multi-drug resistant
mg	milligram
ml	millilitre
MPER	Membrane-proximal external region
mpi	month post infection
mRNA	Messenger ribonucleic acid
MSM	Male sex to male
MVC	Maraviroc
nAb	Neutralizing antibody
Nef	Negative Regulatory Factor
ng	nanogram
ºC	Degrees Celsius
PBMCs	Peripheral blood mononuclear cells
PBS	Primer binding sequence
PCR	Polymerase Chain Reaction
PEG	polyethylene glycol
PEI	Polyethylenimine
PHA-P	Phytohemagglutinin-P
PNGs	Potential N-linked glycosylation sites
pol	Coding for reverse transcriptase
PSVs	Pseuodvirions

RC	Replication capacity
RER	Rough endoplasmic reticulum
rev	Coding for Trans activating protein
RIP	Recombinant Identification Program
RLU	Relative luciferase units
RT	Reverse transcriptase
SDM	Site direct mutagenesis
SDS	Sodium Dodecyl Sulphate
SDS-PAGE	Sodium Dodecyl Sulfate – Polyacrylamide Gel Electrophoresis
SGA	Single genome amplification
SIV	Simian immunodeficiency virus
STIs	Sexually transmitted infections
Tat	Trans-activator of transcription
TBS	Tris-buffered saline
TBS-T	Tris-buffered saline + Tween 20
TCID50	Tissue culture infectivity dose 50 %
TF	Transmitted founder
TMD	Transmembrane domain
UDS	Ultra-deep sequencing
URFs	Unique recombinant forms
USA	United State of America
V1-V5	Variable regions
Vif	Viral infectivity factor
VL	Viral load
Vpr	Viral Protein R
Vpu	Viral Protein U
µg	micrograms
µl	Microliter
5- FOA	5 – flouro – 1, 2, 3, 6 – tetrahydro – 2, 6 – dioxo – 4 – pyrimidine carboxylic acid
6HB	Six helices bundle

ABSTRACT

Identification of HIV-1 Envelope (Env) fitness determinants could provide functionally constrained, accessible regions that could be included in subunit vaccines to induce broadly neutralising antibodies (bnAb). We hypothesised that Env fitness determinants are common to circulating variants but that the plasticity of Env structure limits identification. Rapid evolution; however, could select for sequence changes within the determinants coincident with alterations in function, making identification easier. Dual infection with two phylogenetically distinct HIV-1 variants under the same selective pressures might result in rapid functional evolution, facilitating identification of Env fitness determinants. It has been shown that the Env plays a significant role in viral adaptation to the host environment, which then increases disease progression. Therefore, this study used dual infections as a model system to characterise Env function, its role in *in vivo* viral outgrowth of variants and disease progression and to identify fitness determinants for future vaccine design.

Single-genome amplification (SGA)-derived *env* sequences of four dual infected individuals sampled at enrolment (0 months), 3, 6, and 12 months post infection (mpi) were analysed using Highlighter plots, RIP, DNA pairwise distance and Neighbour-joining trees to determine the *in vivo* evolution of infecting viral populations and their relative frequency over time within each participant. Representative amplicons were cloned at each time point and compared using a pseudovirus (PSV) entry efficiency assay. PSV infection of Affinofile cells induced to carry levels of CCR5 and CD4 similar to T-lymphocytes (CD4high/CCR5low) correlated significantly with that of TZM-bl entry ($p = 0.049$, $r = 0.50$), suggesting that PSV entry of TZM-bl cells was an appropriate model to represent infection of CD4+ T cells.

In three participants recombinant viruses were identified early during infection, suggesting either rapid recombination within the dual infected individual or recombination within the donor prior to transmission. These recombinants had an apparent fitness advantage as they outgrew other viruses at 12 mpi. However, the apparent *in vivo* dominance of recombinants was not associated with enhanced Env entry efficiency for 2/3 individuals, suggesting that

recombination within Env did not always select for fitter variants. The variants that became dominant over time were further characterised by Affinofile system, T-20 IC50 and Western blotting to identify whether tropism, Env expression/cleavage, incorporation into viral particles and fusogenicity were most likely responsible for the variation in Env entry efficiency. All variants were R5- and T-tropic and only Env fusion capacity correlated significantly with Env entry efficiency data ($p = 0.02$, $r = 0.59$), suggesting that variants infecting dual infected participants evolved towards higher fusion capacity. Changes in Env fusogenicity indicated that gp41 might be a fitness determinant of PSV entry efficiency and analysis of SGA sequences indicated that recombination within gp41 was common to 3/4 participants. Env chimeras were generated where gp41 was swapped between clones that either had the same (CAP84) or different (CAP267) PSV entry efficiency. For both participants, and (CAP137) gp41 was identified as a potential determinant of Env fitness. Moreover, two potential N-glycan sites (PNG) at position N332 and N339, previously reported to be involved in neutralising antibody escape, were also identified. While N332 enhanced Env entry efficiency in one participant, N339 attenuated Env entry efficiency in another, potentially due to the escape mutation carrying a fitness cost. However, neither PNG seemed to affect Env expression/cleavage, incorporation into viral particles and fusogenicity.

As Env phenotypic characterisation focussed on PSV assays, we wanted to determine whether viral replication was also similarly affected. Infectious molecular clones (IMCs) were generated from two participants using a recombination yeast assay and replication capacity (RC) in peripheral blood mononuclear cells (PBMCs) was assessed using parallel replication. A significant correlation between RC of viruses in PBMCs and Env entry efficiency in TZM-bl and fusion capacity ($p = 0.03$, $r = 0.7$; $p = 0.04$, $r = 0.7$, respectively) was determined. IMC RC was also associated with *in vivo* outgrowth of viral populations at 12 mpi although this relationship did not always coincide with the frequency of individual variants. Changes in the RC of the Env chimeras and mutants was not associated with phenotypic changes, suggesting that Env entry efficiency determinants did not play the same role in IMC RC as it did in PSV entry.

Lastly, there was a significant negative association ($p = 0.046$, $r = -0.59$) between Env entry efficiency and CD4+ T decline, a marker of disease progression, supporting the previous finding that Env entry efficiency could be the driving agent of disease progression. This was also corroborated by the trend in association between RC of IMCs and faster CD4+ T decline.

Our findings suggest that despite different host pressures, viral competition in most dual infected individuals selected for rapid recombination within gp41 that enhanced fusion capacity. Enhanced gp41 fusogenicity of the dominant viral population at 12 mpi increased PSV entry efficiency and replicative fitness enabling viral outgrowth. Therefore, vaccines that target gp41 might prevent HIV infection or at least attenuate viral fitness and slow disease progression. On the other hand, we showed that targeting the PNG at position N339 of gp120 might influence viral fitness and increase viral load and/or decrease CD4 T cell count. This is in keeping with the association between CD4 T cell decline and PSV entry efficiency and IMC RC, suggesting that Env fitness plays a role in HIV pathogenicity.

CHAPTER 1.

Literature Review

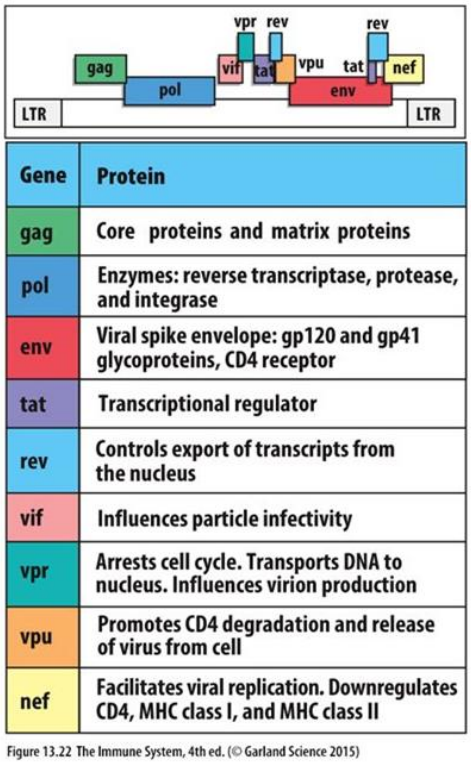
1.1 Introduction

The greatest challenge to vaccine design is the high diversity of Human immunodeficiency virus-1 (HIV-1) as the design of one immunogen based on one population might not elicit a protective immune response to another divergent strain. The rapid introduction of point mutations and homologous swapping of regions by recombination generates a vast array of inter-related variants (Charpentier et al., 2006; Smyth et al., 2012). Therefore, the presence of two different viral strains within a single individual provides an opportunity to generate highly divergent variants within a short period of time (Smyth et al., 2012; Tebit et al., 2007). Whether these variants survive will depend on phenotypic changes that provide a replicative advantage and thus increase in viral fitness. Survival advantages can include resistance to anti-retroviral therapy (ART) (Smith, Wong, et al., 2005), escape from immune responses (van Gils et al., 2010; Song et al., 2012; Troyer et al., 2009) and changes in function of viral proteins (Gordon et al., 2016) all of which can ultimately lead to increased viral load and faster disease progression. This review will explore the importance of dual-infections within the context of viral fitness and potential consequences to HIV-1 pathogenesis.

1.2 Overall HIV structure

HIV-1 genome comprises nine viral genes encoding structural and accessory viral proteins (Figure 1.1) (Frankel and Young, 1998). There are six accessory proteins, namely Vif, Vpr, Tat, Rev, Vpu and Nef, that function as *trans*-activators and regulators of gene expression and three structural proteins, Gag, Pol and Envelope (Env) (Varmus, 1988) essential to the production of infectious particles. HIV-1 Protease, Reverse Transcriptase (RT) and Integrase are encoded by *polymerase* (*pol*) with Protease required for cleavage of viral polyprotein

precursors into mature functional proteins (Johnson and Desrosiers, 2002; Wu, 2004). The polyprotein precursor, Gag-p55 (Mervis et al., 1988; Veronese et al., 1988) is encoded by *gag* and is cleaved into: 1) matrix (MA) (p17) that initiates the budding of viral particles (Ono et al., 2000), 2) capsid (CA)(p24) that encloses the viral genome and is important in assembly of infectious viral particles (Göttlinger et al., 1989), 3) p6, important for the incorporation of the accessory protein, Vpr into budding virus (Kondo and Göttlinger, 1996) and 4) nucleocapsid (p7) that helps to direct viral RNA transport and packaging, efficient reverse transcription and viral infectivity (Levin et al., 2005; Poon et al., 1996).



Gene	Protein
gag	Core proteins and matrix proteins
pol	Enzymes: reverse transcriptase, protease, and integrase
env	Viral spike envelope: gp120 and gp41 glycoproteins, CD4 receptor
tat	Transcriptional regulator
rev	Controls export of transcripts from the nucleus
vif	Influences particle infectivity
vpr	Arrests cell cycle. Transports DNA to nucleus. Influences virion production
vpu	Promotes CD4 degradation and release of virus from cell
nef	Facilitates viral replication. Downregulates CD4, MHC class I, and MHC class II

Figure 13.22 The Immune System, 4th ed. (© Garland Science 2015)

Figure 1.1. HIV-1 genes and proteins. Relative juxtaposition of the nine HIV-1 genes along the genome, the proteins they encode and their functions. http://slideplayer.com/slide/9432845/29/images/26/The+genes+and+proteins+of+HIV-1.jpg

1.3 HIV Diversity

HIV-1 is characterized by extensive genetic diversity driven by many factors: the "*in vivo*" rapid turnover rate of the virus (~ 2.6 days/replication cycle) (Ho et al., 1995), high replication rate of the virus (10^{10} viral particle/day) (Tebit et al., 2007), high mutation rate and lack of efficient proof-reading activity of HIV RT (Mansky and Temin, 1995), recombination within the quasispecies (Charpentier et al., 2006) or between two phylogenetically distinct variants within a single cell in the same individual (Smyth et al., 2012) and the selection pressure of host immune responses [cytotoxic T-lymphocytes (CTL) and neutralizing antibody (nAb)] or ART (Frost et al., 2005; Michael, 1999). The evolution rate was estimated to be 0.0024 and 0.0019 substitutions per base pair (bp) per year for *envelope* (*env)* and *gag*, respectively (Capel et al., 2012). Remarkably, the virus is able to tolerate high mutation rates and maintain its replication fitness (Ndung'u and Weiss, 2012).

Based on phylogenetic analysis, HIV is divided into two genetically distinct types, type 1 (HIV-1) and type 2 (HIV-2), both able to cause AIDS although type 1 accounts for most global HIV infections (Kerina et al., 2013). Each type is then further subdivided into a hierarchy of groups, subtypes (or clades) and sub-subtypes (Figure 1.2). HIV-1 is classified into three groups: major (M), outlier (O), and non-M/non-O (N). HIV-1 Group M is the pandemic branch of HIV-1 and further classified into nine subtypes; designated with letters, A, B, C, D, F, G, H, J and K, distributed globally according to region (Figure 1.3). Subtype C is the most prevalent virus across eastern and southern Africa, contributing 52 % to HIV-1 infections globally (www.unaids.org). The highest genetic diversity between these subtypes is about 30 % within *env* and about 20 % and 15 % within *gag* and *pol*, respectively, (Kerina et al., 2013; Tebit et al., 2007).

Additionally, recombination between different subtypes has given rise to circulating recombinant forms (CRFs) and unique recombinant forms (URFs), that are responsible for 20 % of global infections (Hemelaar et al., 2011) (Figure 1.3). Virus strains are designated as CRFs if they have identical recombination breakpoints (Robertson et al., 2000; Tebit et al., 2007) and are isolated from at least three epidemiologically unlinked individuals, capable of establishing an epidemic on its own (Hemelaar et al., 2011; Kerina et al., 2013). Thus far,

there are 48 different CFRs described (Hemelaar et al., 2011) such as CRF01-A/E responsible for the epidemic in Southeast Asia while CRF02-A/G is predominant in West and West Central Africa, as they represent between 50 % to 70 % of the circulating strains (Figure 1.3) (Peeters, 2000; Requejo, 2006; Tebit et al., 2007). However, URFs have heterogeneous recombinant breakpoints (Tebit et al., 2007) with no evidence of epidemic spread (Kerina et al., 2013).

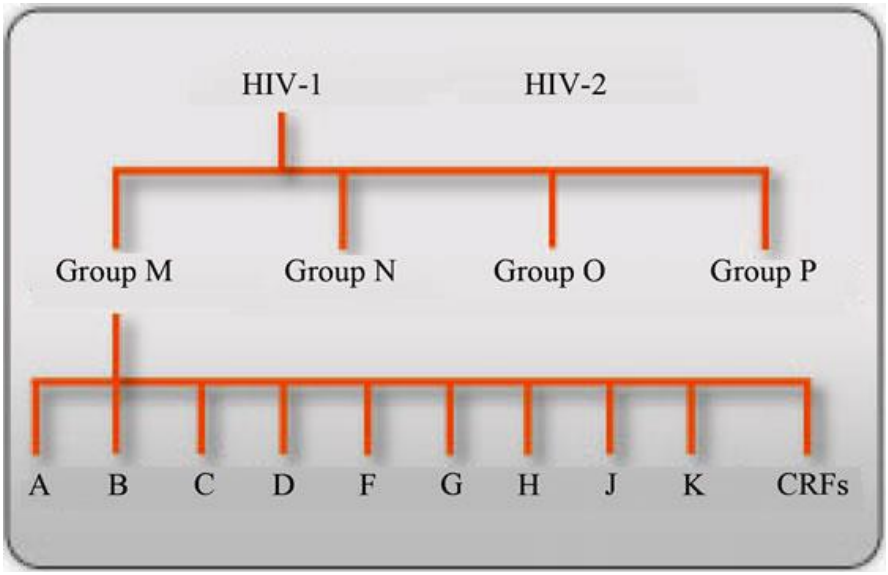


Figure 1.2. HIV Classification. Human immunodeficiency virus is first grouped into type 1 and type 2, with the former driving the global HIV pandemic. HIV-1 is classified into 4 groups: M, N, O and P. Group M is further divided into nine distinct subtypes (A, B, C, D, F, G, H, J and K) as well as circulating recombinant forms (CRFs) (Kerina et al., 2013).

HIV-1 subtypes might influence the efficacy of vaccines as plasma neutralization potency were significantly greater against clade-matched viruses than against mismatched variants (Hraber et al., 2014). Thus, polyvalent mosaic immunogens, assembled by recombination of natural strains of HIV-1, is one promising vaccine strategy to expand cellular immunologic coverage against genetically diverse circulating viruses (Barouch et al., 2010; Santra et al., 2010).

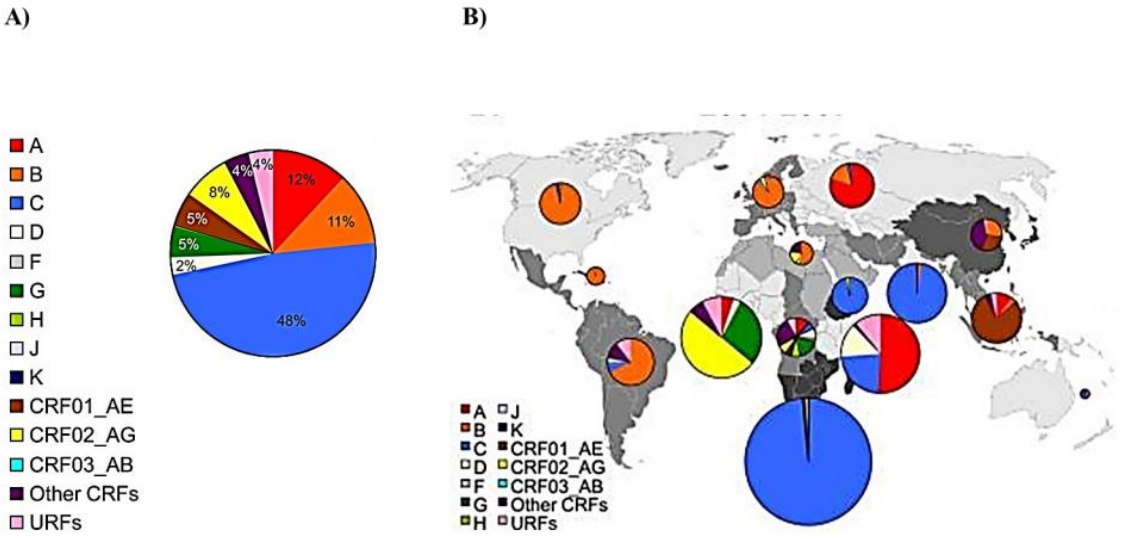


Figure 1.3. Distribution of HIV-1 subtypes worldwide. A) The relative global dominance of HIV-1 subtypes including both circulating recombinant forms (CRFs) and unique recombinant forms (URFs) between 2004-2007. B) The dominant subtype indicated globally where the size of the circle indicates the extent of the pandemic in each region with subtype C driving the epidemic in South Africa. Modified from (Hemelaar et al., 2011).

1.4 Mechanisms and impact of Diversification

HIV-1 is a retrovirus that replicates through DNA intermediates generated by the error prone RT. During viral replication, misincorporation of nucleotides by RT results in a mutation rate of between 1.4×10^{-5} (Abram et al., 2010) to 3.4×10^{-5} mutations per bp per replication cycle (Mansky and Temin, 1995). Given that HIV replicates very rapidly with a life cycle of 1 to 2 days (Ho et al., 1995; Wei et al., 1995) and millions of cells become infected in blood and different lymphoid compartments, the potential for high genetic variation is considerable. Although mutations may occur randomly, mutations that favour effective replication capacity and growth advantages will be selected (Desrosiers, 1999). In addition, mutations that rescue the virus from host immune responses will also provide a selective advantage even if they reduce the replication capacity of the virus (Desrosiers, 1999; Liu et al., 2007).

1.4.1 Point mutation

Point mutations that lead to the shifting, loss or gain of potential N-glycosylation sites (PNGs) are well documented in gp120 and gp41 (Kalia et al., 2003; Wang et al., 2013). Recently a study showed that a PNG shift from position 332 to 334 in Env during chronic infection enabled escape from broadly neutralising antibodies (BCN) (van den Kerkhof et al., 2016; Lynch et al., 2015; Moore et al., 2012). However, escape from antibody neutralisation is not only due to changes in N-glycosylation as point mutations within highly conserved structural motifs such as within the intracytoplasmic tail (CT) of gp41 also resulted in escape from neutralisation due to concomitant conformational changes within gp120 (Kalia et al., 2005).

HLA-associated point mutations have been shown to drive viral evolution in acute HIV-1 infection (Crawford et al., 2009; Goepfert et al., 2008; Goonetilleke et al., 2009; Li et al., 2007; Wang et al., 2009). Mutation within CD8+ cytotoxic T-cell (CTL) epitopes that conferred escape from CTL responses were detected within Gag before seroconversion (Altfeld et al., 2001) and individuals carrying certain HLA alleles such as B*57, B*5801 and B*27 in Gag tended to have slowed disease progression (Brockman et al., 2007; Chopera et al., 2008; Crawford et al., 2009; Leslie et al., 2004; Martinez-Picado et al., 2006). This was because CTL escape mutations within epitopes that fell within conserved regions had a detrimental effect on Gag function (Leslie et al., 2004; Martinez-Picado et al., 2006) leading to decreased replication fitness (Liu et al., 2007; Sunshine et al., 2015; Troyer et al., 2009). Previously, Miura et al., (2008) reported that in elite controllers (EC), selection of rare escape variants by HLA-B*57 resulted in severely compromised viral replication (Miura et al., 2008).

Additionally, point mutations also confer drug resistance to RT and protease inhibitors that have been linked to lowered viral fitness (Nijhuis et al., 2001; Quiñones-Mateu and Arts, 2006). CCR5 inhibitors such as maraviroc (MVC) and vicriviroc (VVC) bind to CCR5 preventing viral entry (Kuhmann and Hartley, 2008). However, MVC-resistance can be conferred by mutations within the CD4-binding site of Env while VVC-resistant variants are able to bind to inhibitor-bound CCR5 due to point mutations within the V3 loop and gp41 fusion peptide (Anastassopoulou et al., 2009; Pugach et al., 2007; Tsibris et al., 2008)

although these latter changes were not linked to viral fitness (Anastassopoulou et al., 2007). Amino acid changes in gp41 associated with resistance to enfuvirtide (ENF/T20) has been associated with decreased replicative fitness (Lu et al., 2004; Reeves et al., 2005) although this was not always the case (Neumann et al., 2005). Therefore, single or multiple point mutations not only allow escape from immune responses and resistance to drug therapy, it can also lead to concomitant changes in viral fitness which might then influence disease progression.

1.4.2 HIV-1 Recombination

HIV-1 is a diploid virus with two RNA genomes packed into one virion. During virus replication, RT may jump from one RNA template to another, generating genomic sequences comprising two (or more) parent genomes (Figure 1.4). If the two parent viruses have non-identical genetic information, then the newly packaged RNA molecules will be a mosaic virus that might have "novel" properties (Blackard et al., 2002; Charpentier et al., 2006; Jetzt et al., 2000). These new "heterogeneous" virions continue to replicate and form a quasispecies of recombinant variants. Recombination was mainly found in gag/*pol* and *env* (Charpentier et al., 2006; Santoro and Perno, 2013; Simon-Loriere et al., 2009) and events can occur from 2-20 per genome per replication cycle (Charpentier et al., 2006). Both inter-subtype (different HIV subtypes) (Magiorkinis, 2003), and intra-subtype (same subtype) (Kiwelu et al., 2013; Rousseau et al., 2007) recombination between two or more phylogenetically distinct viruses have been reported.

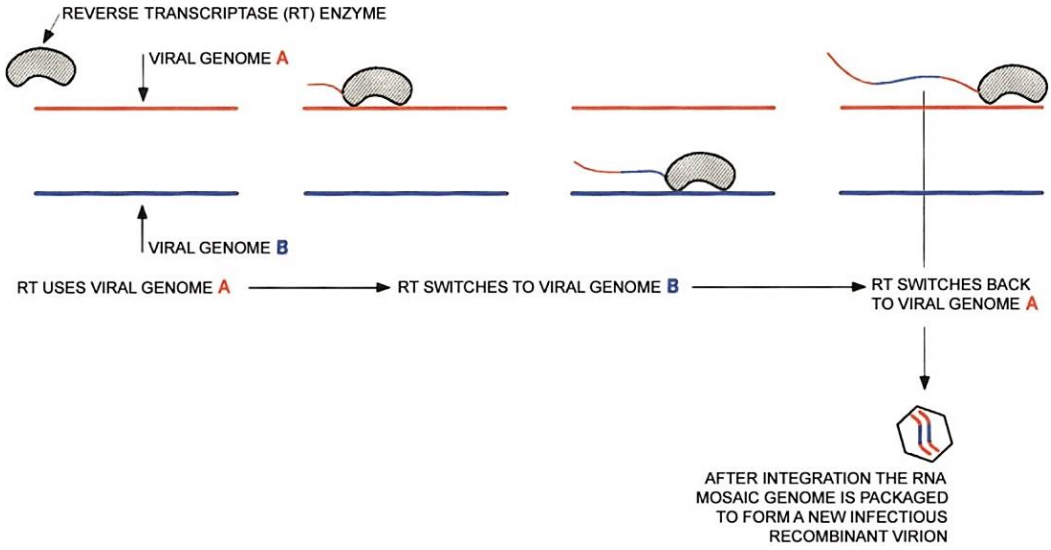


Figure 1.4. Template switching during HIV-1 reverse transcription. In HIV and other retroviruses, the reverse transcriptase (RT) enzyme can shuffle from one virus template to another during the formation of the double stranded DNA provirus. This generates new mosaic infectious virions carrying a mixture of parent viral genomes. Figure was adapted from Blackard et al., (2002).

Similar to point mutations, recombination may also confer variants with advantageous traits, such as escape from immune recognition and drug resistance, resulting in enhanced pathogenicity (Mostowy et al., 2011). Furthermore, recombination can change viral tropism and virus replicative fitness and thus accelerate disease progression (Figure 1.5) (Blackard et al., 2002; Mostowy et al., 2011). Recombinants can also be transmitted, resulting in the spread of highly fit variants and altering the natural history of the HIV-1 epidemic (van der Kuyl and Cornelissen, 2007).

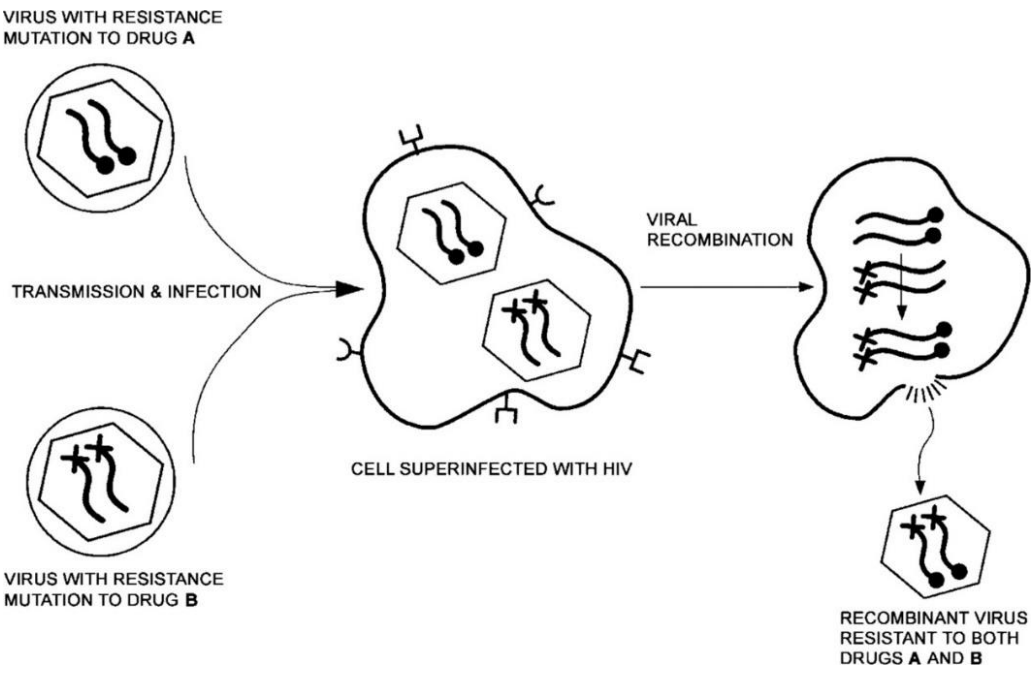


Figure 1.5. Recombination between 2 HIV viruses can select for fitter viruses. Recombination between two HIV virions with differing drug resistance profiles can generate a new infectious recombinant virus resistant to both drugs. Figure adapted from Blackard et al. (2002).

1.5 HIV-1 transmission

The majority of HIV-1 transmission cases occur heterosexually while only 20 % of infections are due to percutaneous or intravenous inoculations (Cohen et al., 2011). The genital tract mucosa forms a protective barrier against HIV infection. However, many factors can disrupt this barrier such as inflammation, infection of vaginal epithelium as in the case of sexually transmitted infections (STIs), and hormonal contraceptives (Sagar et al., 2004), increasing the risk of HIV infection. Following transmission, there is a virus bottleneck. Despite the high diversity of HIV quasispecies in donors, a homogenous viral population was identified in newly infected individuals (Derdeyn et al., 2004). Studies on heterosexual transmission estimated that approximately 80 % of infections was as a result of transmission of a single

virus termed the transmitted founder (TF) (Abrahams et al., 2013; Keele et al., 2008) while only 60 % of men who have sex with men (MSM) and about 40 % of injection-drug users (IDU) became infected with a single virus (Cohen et al., 2011). Subtype C and A TF Envs have distinct phenotypes such as CCR5 coreceptor tropism, shorter V1-V2 loop lengths that are less N glycosylated, and greater sensitivity to antibody neutralization (Chohan et al., 2005; Derdeyn et al., 2004). Other studies reported the transmission of multiple variants (quasispecies) (Grobler et al., 2004; Sagar et al., 2003; Woodman et al., 2011) where more than one virus, transmitted from the donor to the recipient, established infection. The frequency of multiple variant transmission varied greatly between these studies most likely influenced by virus subtype, study population and the route of infection (Haaland et al., 2009). However, even in the case of multiple variants, there is evidence to suggest that the transmission bottleneck selects for virus with higher replication fitness (Joseph et al., 2015) although, this is not always the case (Song et al., 2016).

Two scenarios have been proposed to explain the transmission bottleneck (Figure 1.6) and how it results in the transmission of a single virus: the first scenario suggests that after transmission of multiple variants only the fittest virus survives and establishes clinical infection while viruses with poor fitness are lost. In the second scenario, after exposure to HIV-1, only one virus replicates at the site of transmission and establishes systemic infection (Cohen et al., 2011; Joseph et al., 2015).

However, the mechanism of multiple variant transmission is still unclear. Recent studies hypothesised that genital tract infections and STIs may destroy the intact mucosa and increase the availability of activated CD4+ T cells, thus increasing the rate of multivariant transmission (Haaland et al., 2009; Woodman et al., 2011). Alternatively, transmission of multiple variants are linked events: the transmission of one virus enhances the transmission of a second virus (Abrahams et al., 2013). It has been shown that high risk individuals are more likely infected with recombinants or phylogenetically distinct variants, suggesting that multiple sexual partners might provide a source of URF and fitter variants that contribute to increased viral loads and lowered CD4 counts (Herbinger et al., 2006; Ssemwanga et al., 2011).

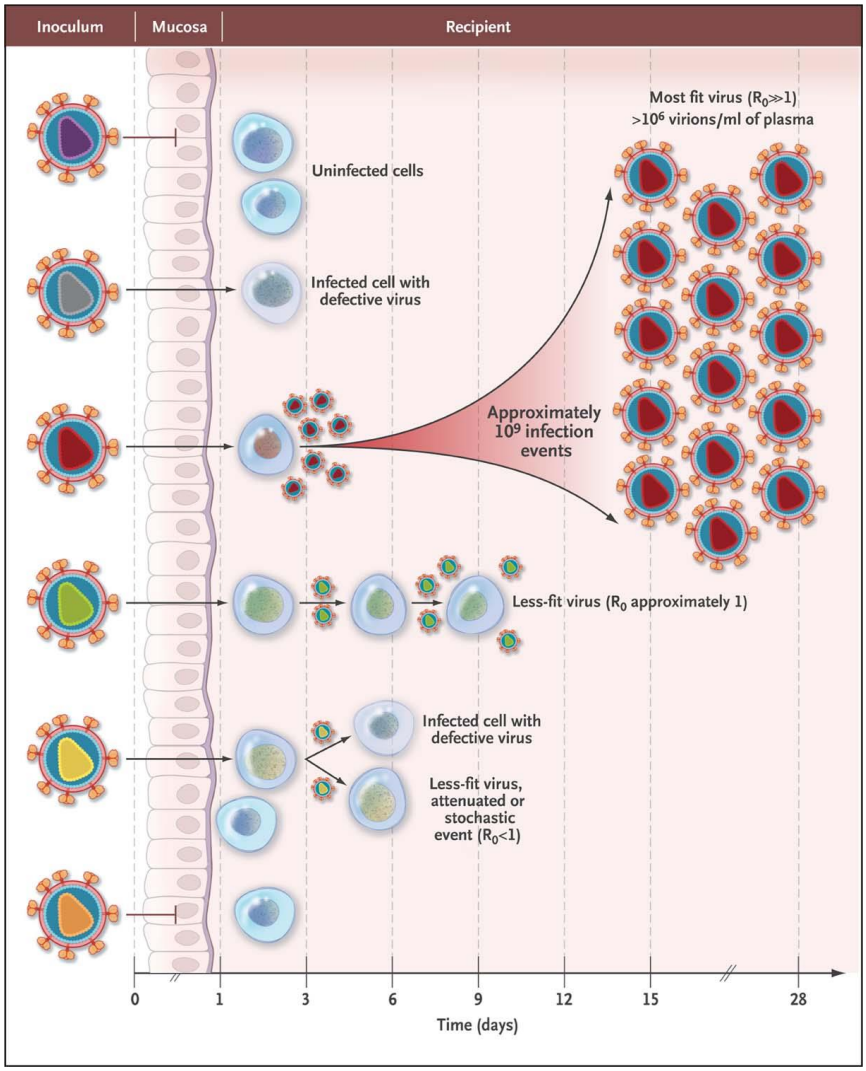


Figure 1.6. HIV-1 transmission bottleneck. Different scenarios proposed during transmission in which a single "fit virus" is able to establish productive infection while infection with a less fit variant or defective virus cannot establish an infection even after crossing the mucosal barrier. Ro indicates the reproductive ratio of the transmitted virus, where Ro <1 indicates terminated infection. Taken from Cohen et al., (2011).

This thesis examines and characterise viruses isolated from four high-risk women from the CAPRISA 002 cohort in South Africa. In a previous study (Woodman et al., 2011) these participants were identified to be infected with two phylogenetically distinct variants at

enrolment, suggesting the transmission of more than one variant during early stages of infection either simultaneously or within a short time before seroconversion. Moreover, these participants were diagnosed to have one or more STIs at enrolment which most likely contributed to the disruption of the mucosal barrier and thus increased the susceptibility to HIV infection (Woodman et al., 2011). It has been hypothesised that coinfection of highly diverse variants would provide the sequence space for rapid selection of fitter variants leading to enhanced disease progression. However, there have been conflicting evidence as to whether multivariant transmission is associated with increased disease progression (Andreani et al., 2011; Grobler et al., 2004; Sagar et al., 2003; Woodman et al., 2011).

1.6 HIV-1 disease progression

Advanced techniques enable researchers to detect viral RNA and early immune responses within the first few weeks of HIV-1 infection (McMichael et al., 2010). This period is associated with impaired early innate immune responses and damage to the genital mucosa due to viral cytopathicity (McMichael et al., 2010).

Viral RNA can be detected 7 to 21 days after HIV-1 transmission and this period is known as the "eclipse phase" and refers to the time from virus entry into a cell to production of new virions (Figure 1.7) (Cohen et al., 2011; Haase, 2011; McMichael et al., 2010). During this time, virus replicates exponentially in the mucosa, submucosa and draining lymphoreticular tissues, targeting different cell types including Langerhans' cells (LCs), dendritic cells (DCs), resting CD4+ T cells (Haase, 2011) and activated CD4+ T cells (Cohen et al., 2011). In addition, the virus establishes latent infection in long lived resting CD4 cells (Haase, 2010).

Following the eclipse phase, the virus and/or virus-infected cells spread to other lymphoid tissues and the gut-associated lymphoid tissue (GALT) where activated CD4+CCR5+ memory T cells are found in abundance for rapid replication (McMichael et al., 2010). Plasma viremia levels can peak at 10^7 or more copies of viral RNA per millilitre of blood (Coffin and Swanstrom, 2013) at around 21-28 days (Haase, 2011; McMichael et al., 2010). During this period known as peak viremia, the number of CD4+ T cells are significantly

depleted and acute phase symptoms might appear (Haase, 2011) (Figure 1.7). The first immune response modulated by DCs and natural killer (NK) cells appear during this stage and is characterised by a large burst of inflammatory cytokines (Cohen et al., 2011). In addition, CD8+ T cell responses and non-neutralising anti-HIV antibodies (seroconversion) have also been detected at peak viremia (Gaines et al., 1988). Tomaras et al. (2008) reported the first B cell response in the form of plasma immune complexes 8 days after HIV detection and the first anti-HIV-1 antibody 13 days after the appearance of virus. Early events are critical to HIV survival and this period is considered a window of opportunity for vaccine or early intervention to prevent systemic infection (Fauci, 2007; Haase, 2011).

After peak viremia, the viral load decreases and stabilises over 21-20 weeks known as viral set point. At the same time, CD4+ T cell numbers increase to near normal concentrations in blood although counts continue to fall in the GALT. However, blood CD4+ T cell numbers begin to decline over time and is a surrogate marker for disease progression (Maartens et al., 2014). The chronic phase of infection (known also as clinical latency), can extend from 1-20 years and infected individuals can remain asymptomatic throughout (Coffin and Swanstrom, 2013). This stage is maintained by a balance between virus turnover and immune responses that limit viral replication (McMichael et al., 2010) until viremia increases and the number of CD4+ T cells declines to the point at which the immune system cannot control adventitious infectious agents (Coffin and Swanstrom, 2013).

Based on viral loads and CD4+ T cell counts, the clinical course of HIV infection was classified into three stages: 1) primary infection or acute phase of infection, when the viral load is high, 2) clinical latency, when the viral load is stable, and 3) AIDS-defining stage, where the viral load is high and host immunity is destroyed (Figure 1.7) (Mellors, 1997; Mylonakis et al., 2001). However, the rate of disease progression is highly variable between infected individuals and they are categorised as rapid, typical (or intermediate), slow or long-term non progressors (LTNPs) (Langford et al., 2007; Pantaleo and Fauci, 1996) and EC (Zaunders and van Bockel, 2013).

Most HIV infected individuals (70-80 %) are typical progressors that follow the clinical course of infection described above. Typical progressors have long asymptomatic periods of

clinical latency (8-10 years) with CD4+ T cell counts usually higher than 500 cells/µl (Jurriaans et al., 1994; Pantaleo and Fauci, 1996). Progression to AIDS-defining illness occurs when the CD4+ T cell counts drops and plasma viral load increases (Langford et al., 2007; Mellors, 1997; Pantaleo and Fauci, 1996; Phillips and Pezzotti, 2004). Initially, ART was introduced once CD4+ T cell levels dropped to 200 cell/ul but recently it has been debated whether ART should be initiated soon after HIV-1 infection irrespective of CD4+ levels (Ying et al., 2016). Rapid progressors on the other hand, comprising 10-15 % of infected individuals, progress to AIDS within two to three years. These individuals experience prolonged acute infection with no clinical latency transition. The CD4+ T cell counts decline very rapidly to less than 350 cell/µl within the first year of infection (Mlisana et al., 2014) with a rapid rise in viral load (between 3000 copies/mL to ≥ 300 000 copies/mL) (Langford et al., 2007). However, in LTNP, the viral load drops after acute infection and their CD4+ T cell counts remain within the normal range for several years (eight to ten) (Mandalia et al., 2012; Pantaleo and Fauci, 1996; Zaunders and van Bockel, 2013) whereas ECs suppress viral load to below detectable levels (Mandalia et al., 2012).

Variation in disease progression of HIV-1 infected individuals could be due, in part, to the genetic complexity/diversity of infecting viruses impacting the rate of disease progression (Sagar et al., 2003). Another study showed that diversification over the course of infection also influenced the rate of disease progression (Mani et al., 2002). Thus, the genetic diversity of viral populations circulating in infected individuals may play a role in determining the severity of the disease.

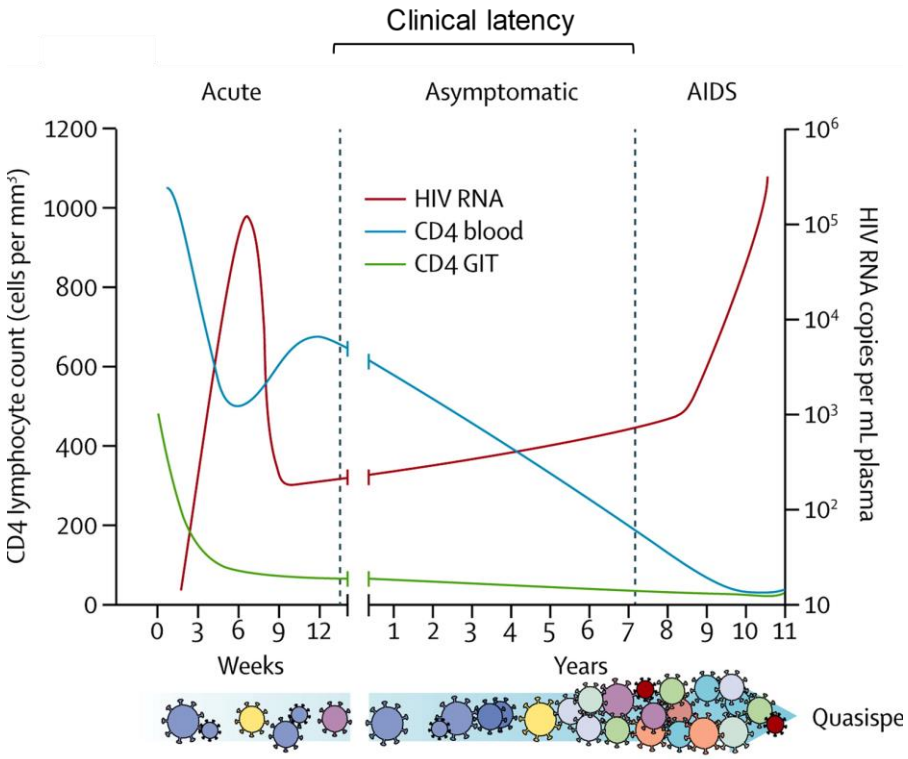


Figure 1.7. Time course of HIV-1 infection and disease progression. Average HIV-1 disease progression of untreated typical progressors. CD4+ T cells are depleted progressively in the gastrointestinal tract (GIT) (green line) and remain low throughout infection while CD4+ T cells in blood (blue line) are depleted during acute phase of infection but rise again during the asymptomatic period also called "clinical latency". HIV RNA copies (viral load) (red line) peak during acute infection and stabilise during clinical latency before rapid viral replication gives rise to a highly diverse viral population (quasispecies) when CD4+ T cells are depleted, and AIDS-defining illness appear. Graph adapted from (Maartens et al., 2014).

1.7 Dual infection

Grobler et al., (2004) defined dual infection with the same subtype "as an infection with 2 phylogenetically distinct viruses that are no more closely related to each other than are another pair of epidemiologically unlinked viruses, with a mean pairwise DNA distance at least as distant as that to a group of unlinked sequences". Dual infection is classified into two groups based on the timing of infection: 1) coinfection of two viral strains occurs prior to seroconversion, either simultaneously or within a brief period of time (Jobes et al., 2006),

(this usually takes anywhere from a few weeks to a few months), and 2) super-infection, also known as reinfection, when an HIV-positive individual becomes infected with a second strain after seroconversion (Gottlieb et al., 2004; van der Kuyl and Cornelissen, 2007). Super-infection has been the focus of several studies due to its relevance to vaccine design: infection of one virus does not protect against infection by others. However, coinfection with multiple strains prior to adaptive immune responses might also be relevant to the efficacy of vaccine trials (Jobes et al., 2006) and the design of multiple subunit/epitope vaccines (Powell et al., 2010). Phylogenetically distinct strains can be transmitted from the same donor or from multiple partners (Ssemwanga et al., 2011) and dual infection can occur between CRFs and URFs (Girard et al., 2006; Smith, Richman, et al., 2005) or within subtypes (Artenstein et al., 1995), between HIV-1 and HIV-2 (Georgoulias, 1988) as well as between HIV-1 groups (De Oliveira et al., 2017).

1.7.1 Incidence of dual infection

The prevalence of CRFs and URFs globally suggests high frequency of dual infections (Girard et al., 2006), although identification of dual infections is rare (Smith, Richman, et al., 2005) with none detected in studies of chronically infected individuals in more than 1072 and 215 person-years of study (Gonzales et al., 2003; Tsui et al., 2004). Other studies, however, have identified HIV dual infection during chronic infection, (Brenner et al., 2004, p. 1; Fang et al., 2004; Grobler et al., 2004; Pernas et al., 2006), suggesting that susceptibility to infection with more than one strain is not restricted to early HIV infection.

The earliest report of dual infection, published in 1987, was on an experimental animal model. The aim of the study was to infect chimpanzees with two different strains of HIV-1. After the first infection, no AIDS-like syndrome was observed in chimpanzees however, after infection with the second strain, high antibody titres to both virus strains were detected (Fultz et al., 1987). In 1994, Sala et al., (1994) reported the first human dual infection when he isolated two distinct viral strains from LCs of skin (Sala et al., 1994). Later, in 1995, two further reports were documented: The first one identified a patient coinfected with multiple strains of subtype B virus (Zhu et al., 1995), and the other reported two patients in Thailand who were infected with subtypes B and E variants (Artenstein et al., 1995). In 1996, Janini

(1996) reported a case of dual infection during the analysis of 33 samples from Brazilian patients, with two distinct HIV-1 subtypes F and D using restriction fragment length polymorphism (RFLP). After that, many reports of dual infection (coinfection and superinfection) were documented (Table 1.1). In 2002, two cases of superinfection were found in Thailand (Ramos et al., 2002) and another case was reported in which a man infected with HIV-1 subtype AE became infected with an HIV-1 subtype B virus (Jost et al., 2002). Both Smith et al., (2004) and Yerly et al., (2004) indicated approximately 5 % incidence of dual infection in IDUs in Geneva and Lausanne, sites of a Swiss Cohort (Smith et al., 2004; Yerly et al., 2004). In 2004, Gottlieb et al, (2004) analysed sixty-four patients from two different cohorts, the Seattle Primary Infection cohort and the South African female sex workers cohort. He found five patients (~8 %) with intra-subtype HIV-1 dual infection, four participants were co-infected with two strains of subtype B, and one was superinfected with a subtype C variant <3 years' post-infection with the first virus (Gottlieb et al, 2004). The frequency of dual infections has been reported to range from 6 to 56 % (Table 1.1). The global variation in the incidence of dual-infections could be due to differences between cohorts, intra-subtype infections, timing of infection and the inability to detect variants at low frequency or intra-subtype recombination (Burke, 1997; Pacold et al., 2010).

Table 1.1 Incidence of dual infections

Type of Dual Infection	Type of recombination	Incidence %	Cohort	Country	Reference
Coinfection and Superinfection	Intra-subtype	8	Seattle Primary Infection Cohort and South African female sex workers Cohort	United states (USA) & South Africa (SA)	(Gottlieb et al., 2004)
Coinfection	Intra-subtype	19	Female sex worker, Durban cohort	SA	(Grobler et al., 2004)
Multiple infection	Inter-subtype	19	High risk female bars cohort	Tanzania	(Herbinger et al., 2006)
Coinfection and Superinfection	Inter-subtype and Intra-subtype	43	Amsterdam cohort (1996-2005)	Netherland	(Cornelissen et al., 2007)
Dual infection	Inter-subtype and Intra-subtype	32	Amsterdam cohort (2003-2007)	Netherland	(van der Kuyl and Cornelissen, 2007)
Coinfection and Superinfection	Inter-subtype and Intra-subtype	43.5 56.5	Women's Interagency HIV Study (WIHS)	USA	Templeton et al. 2009
Coinfection	intra-subtype	9	CAPRISA cohort 002 (2006-2009)	SA	(Woodman et al., 2011)
Coinfection	inter-subtype	21.7	Buenos Aires cohort	Argentina	(Andreani et al., 2011)
Coinfection and Superinfection	Intra-subtype	6 8	San Diego Primary Infection high-risk MSM Cohort	USA	(Wagner et al., 2014)
Coinfection and superinfection	Inter-subtype and Intra-subtype	13.1 15.6	Chinese MSM Cohort	China	(Luan et al., 2017)

1.7.2 Method of Identification

Identification of dual infection requires molecular evidence of the co-circulation of two or more viral populations with a DNA distance greater than within-host evolution from one single founder virus (Pacold et al., 2010). Due to the large genetic differences between subtypes, inter-subtype dual infection is easier to detect than intra-subtype dual infection because of similarity between infecting strains (Burke, 1997; Pacold et al., 2010). Intra-subtype dual infection is limited to molecular and serological methods (Jobes et al., 2006).

Challenges facing the detection and estimation of the incidence of intra-subtype dual infection are multifactorial: insufficient longitudinal sampling, insensitive detection methods to detect viral strains at low frequency, and the rapid emergence of recombinant viruses. Recombination limits the ability to detect initial parent virus based on a single genome fragment and thus analysis of partial gene or sub-genomic sequences might lead to an underestimation of the number of HIV-1 dual infected individuals (Jobes et al., 2006; Pacold et al., 2010).

In order to overcome sampling limitations, individuals should be enrolled into studies within weeks of infection and followed longitudinally for years (van Loggerenberg et al., 2008). Moreover, detection of dual infection can be strengthened by applying more than one methodology such as heteroduplex mobility assay (HMA) and single genome amplification (SGA) (Woodman et al., 2011). HMA is one of the earliest methods for identifying dual infection and is considered a reliable method in estimating HIV-1 diversity (Sahni et al., 2007). It can detect minor variants present at low frequency in a background of distinct quasispecies (Delwart et al., 1994). However, the method has limitations such that it probes only a fraction of the gene of interest and not the whole gene or genome (Salazar-Gonzalez et al., 2008). SGA is based on diluting the cDNA template to an endpoint in which ≤ 30 % of subsequent PCR reactions are positive, statistically supporting the notion that each amplicon was derived from a single cDNA template (Keele et al., 2008; Salazar-Gonzalez et al., 2008). This strategy avoids artefacts induced by recombination between multiple templates during PCR (Etemad et al., 2015) and diversity estimates were considered to be a true reflection of the viral population. However, this method also has limitations, as it is

subjected to sample bias and might also underestimate the number of viruses present in a sample (Pacold et al., 2010; Woodman et al., 2011). Next-generation or ultra-deep sequencing (UDS) technologies, providing high resolution estimates, seems particularly effective for monitoring diversity of rapidly mutating viruses such as HIV-1 (Bushman et al., 2008; Eriksson et al., 2008). Pacold et al., (2010) showed that UDS was equally or more effective than SGA at detecting dual infections, able to detect variants at a frequency of 1 % (Pacold et al., 2010). UDS has been used in recent studies to show the prevalence of coinfections and superinfection incidence (Luan et al., 2017; Wagner et al., 2014) suggesting it might be the method of choice to identify dual infections in future studies. However, irrespective of method used, lack of longitudinal sampling and targeting of single genomic regions are likely to lead to underestimating the frequency of dual infections (Luan et al., 2017). The challenge to detect phylogenetically distinct viruses will impact the ability to fully characterize the phenotype of variants infecting dual infected individuals.

1.7.3 Association between dual infection and increased disease progression

Most of the reports available to date are mainly on inter-subtype recombination and superinfection and there are very few reports on the effect of intra-subtype recombination before seroconversion between co-infected phylogenetically distinct HIV strains and their *in vivo* effect on disease progression. It is possible that the timing of recombination and early emergence of fitter viruses within the context of host-specific immune responses could determine how rapidly the infected individual progresses to AIDS-defining illnesses. Therefore, there is a need for longitudinal studies of co-infected individuals to understand how viral fitness could contribute to increased disease progression.

It has been suggested that dual infections influence HIV disease progression, as it is accompanied by a rise in plasma viral load and a decline in CD4+ T cell numbers (Gottlieb et al., 2004; Grobler et al., 2004; Sagar et al., 2003). One of the first studies that showed how recombination between two strains resulted in enhanced pathogenicity involved an *in vitro* study of infected rhesus monkeys. When two attenuated non-pathogenic phylogenetically distinct strains of simian immunodeficiency virus (SIV) with deleterious mutations (delta-*vpx/vpr* and delta-*nef* genes) were used to infect rhesus monkeys simultaneously, a full-

length, proviral DNA of SIVmac239 virus was isolated from PBMCs within two weeks of infection. Furthermore, there was a concomitant increase in viral load and decline in CD4+ T lymphocyte concentrations (Wooley et al., 1997).

A number of human studies also reported similar associations between dual infections and disease progression: a case of MSM transmission where infection of two phylogenetically unlinked viruses was associated with rapid progression to AIDS defining illnesses in less than two years (Liu et al., 1997). Recombination selected for a mosaic virus with increased fitness, suggesting that recombinants that emerge from dual infection accelerated disease progression in this participant (Liu et al., 2002). Subsequent studies also reported that the emergence of recombinant viruses from inter-subtype and intra-subtype recombination was associated with increased viral load and decreased CD4+ T count (Fang et al., 2004; Gottlieb et al., 2004; Grobler et al., 2004; Nájera et al., 2002; Ramirez et al., 2008).

A male volunteer from the world's first Phase 3 HIV vaccine efficacy trial (VAX004) got coinfected with two strains of subtype B despite receiving three immunizations. Recombination between the two strains resulted in a highly diverse viral population and rapid disease progression, leading to the commencement of ART (Jobes et al., 2006). Importantly, a recent finding on a dual-infected study participant showed that a recombinant virus emerged within 18 mpi with enhanced viral fitness compared to the parent viruses. This suggested that recombination facilitated the rapid generation of fitter viruses that could drive increased viral load and disease progression (Gordon et al., 2016).

Sagar at al. (2003) reported on 89 women from Kenya, who were infected with multiple variants that lead to rapid disease progression with significantly high viral loads (4.84 log10 copies/ml) and low CD4+ T-cell counts (median 416 cells/ µl) within the first year of infection. In two longitudinal studies on female dual infected sex workers, CD4+ T cell counts dropped to below 200 cells/µl within three years post-infection (Gottlieb et al., 2004; Grobler et al., 2004). Moreover, a case of intra-subtype superinfection with a "wild type" virus after primary infection with a drug resistant virus within the first four months resulted in a sharp increase in viral load with concomitant decline in CD4+ cell counts, suggesting that the second virus had high *in vivo* viral fitness compared to the first virus leading to rapid

disease progression. Sequence analysis showed that this second viral strain had additional polymorphisms (K20R, M36I, L63P) in the protease gene that were not present in the first strain, suggesting that these changes could play a role in the higher *in vivo* fitness of the second strain (Koelsch et al., 2003). Recently, UDS analysis of sixty-four HIV-1 early-infected participants showed that both coinfections and dual infections were associated with increased disease progression (Luan et al., 2017).

On the other hand, other studies have not shown a significant link between dual infections and disease progression. In one study, two individuals were found to be infected with two distinct strains of HIV-1 subtype B and were defined as LTNP and remained asymptomatic without ART for > 15 years (Casado et al., 2007). Another study, suggested that dual infections were not always associated with increased disease progression, but that disease outcome dependended on the interplay between viral and host factors (Woodman et al., 2011). Although a number of reports provide evidence of a significant association between dual infection and disease progression, further studies are still required to identify the extent of the association and factors that influence this relationship.

1.7.4 Dual infection and Drug Resistance

Besides increased disease progression, dual infections are also linked to drug resistance mutations (Figure 1.2). In one study, a lymphoblastoid cell culture was coinfected with a mixture of two distinct viruses. One virus was highly resistant to a protease inhibitor (SC-52151) and the other highly resistant to zidovudine. The recombinant virus isolated was resistant to both drugs (Moutouh et al., 1996). After initial infection with a drug sensitive virus, superinfection with an intra-subtype, drug resistant virus was associated with high viral load and low CD4+ T counts, suggesting enhanced *in vivo* fitness (Smith et al., (2005). Another study of a dual infection between HIV-1 (CFR-02 A/G) and HIV-2 (clade A) (HIV-1/2) was reported in a young heterosexual man. Both viruses conferred multiclass drug resistance mutations that was associated with rapid disease progression (Castro et al., 2012). Recently, one case of dual infection (intra-subtype) with multidrug resistant viruses reported a participant superinfected with a highly divergent multi-drug resistant strain that lead to ART failure and subsequent change in highly active antiretroviral therapy (HAART) (Martin

et al., 2016). Most of these cases refer to superinfection, highlighting the lack of studies on HIV-1 coinfections. However, it is reasonable to suggest that coinfection with drug resistant phylogenetically distinct viruses would also lead to treatment failure and/or enhanced viral fitness.

1.8 HIV-1 Envelope

The first step in HIV-1 infection begins with virus adhesion to the host cell surface via CD4 and CCR5 (and other coreceptors) receptors followed by fusion between viral and host cell membranes and subsequent delivery of the virus core material to the cell cytoplasm. This process is mediated by the viral Env.

1.8.1 Structure and function

HIV-1 Env is a heterodimer of two glycoprotein subunits, the surface glycoprotein gp120 and transmembrane (TM) gp41 that are non-covalently connected to form a trimer on the surface of the HIV virion. Env is co-translationally modified in the rough endoplasmic reticulum (RER) as a gp160 precursor of approximately 845–870-amino acids (Staropoli et al., 2000; Wyatt, 1998), with N glycosylation with high mannose carbohydrates (Freed and Martin, 1995a). Oligomerization of gp160 monomers also occurs in the RER (Checkley et al., 2011). Subsequently, and during transport of the Env precursor to the cell surface, the protein undergoes trimerization and modification of high-mannose sugar chains to form hybrid and complex-type N-glycans (Freed and Martin, 1995a; Li et al., 2000). Lastly, gp160 is cleaved within the trans-Golgi at a K/R-X-K/R-R motif by a cellular protease, Furin generating mature gp120 and gp41 subunits. The cleavage process is critical for viral fusion and infectivity (Blay et al., 2007; Freed et al., 1989) and the trimers are incorporated into the budding virion and displayed as a spike on the surface of viral particles (~ 10–14 spikes/virion) (Didigu and Doms, 2012; Staropoli et al., 2000; Wyatt, 1998). The three gp120 and three gp41 subunits form a cage-like structure as was described by Mao et al., (2012) using ~11-Å cryo-electron microscopy (Mao et al., 2012). However, it has also been shown that uncleaved gp160 as well as monomers are also incorporated into the virus and these

could act as decoys for the immune response as they lack the native, functional structure of Env (Herrera et al., 2005; Moore et al., 2006).

Gp120 consists of five highly conserved regions (C1-C5) that are predominantly located in the inner domain of the protein, or the core, and five highly variable regions (V1-V5) that are located on the outer domain of the protein (Figure 1.8) (Modrow et al., 1987). The two domains are connected by a third domain called the bridging sheet (Koning et al., 2002). The first four variable regions are demarcated by disulfide bridges at each end and are important mainly in modifying the immunogenicity and antigenicity of gp120 due to their high number of N-linked glycosylation sites. In addition, V3 is important for coreceptor binding, viral tropism (Hwang et al., 1991) and membrane fusion (Freed et al., 1991). The conserved regions play a crucial role in gp120 folding and binding to CD4 (Didigu and Doms, 2012; Kwong et al., 1998; Wilen et al., 2012a). Gp41 is about 345 amino acids that are divided into three major domains: an ectodomain (or extracellular), a transmembrane domain (TMD), and a C- terminal cytoplasmic tail (CT) (Figure 1.8) (Montero et al., 2008).

The ectodomain contains: 1) an N-terminal, hydrophobic region known as the fusion peptide (FP) that inserts into the target cell membrane during the fusion/entry process, 2) two heptad repeat regions (HR1 and HR2) which are also known as N-helix and C-helix, respectively, connected by a disulfide bridge within a hydrophilic loop, and 3) the membrane-proximal external region (MPER) that is found to be a highly conserved and Tryptophan-rich (Didigu and Doms, 2012; Montero et al., 2008). The TMD, that anchors the Env in the viral membrane, consists of approximately 25 highly conserved amino acids. It was found to play a role in Env-mediated membrane fusion (Checkley et al., 2011).

The CT consists of different motifs: the internalization signal, YSPL; the Kennedy sequence (ks); the conserved amphipathic α-helices LLP-1, LLP-2, and LLP-3; and a C-terminal dileucine motif (LL) that are involved in endocytosis and intracellular distribution of Env (Checkley et al., 2011; Yue et al., 2009). LLP-1 and LLP-2 are found in the central region and, at the C-terminal region, respectively of gp41 CT, while LLP-3 is located between the other two helices (Figure 1.8A). The first two LLP segments are positively charged regions due to the presence of Arginine residues on one face of the α-helix. These regions are

involved in Env incorporation, cell surface expression and Env fusogenicity (Checkley et al., 2011).

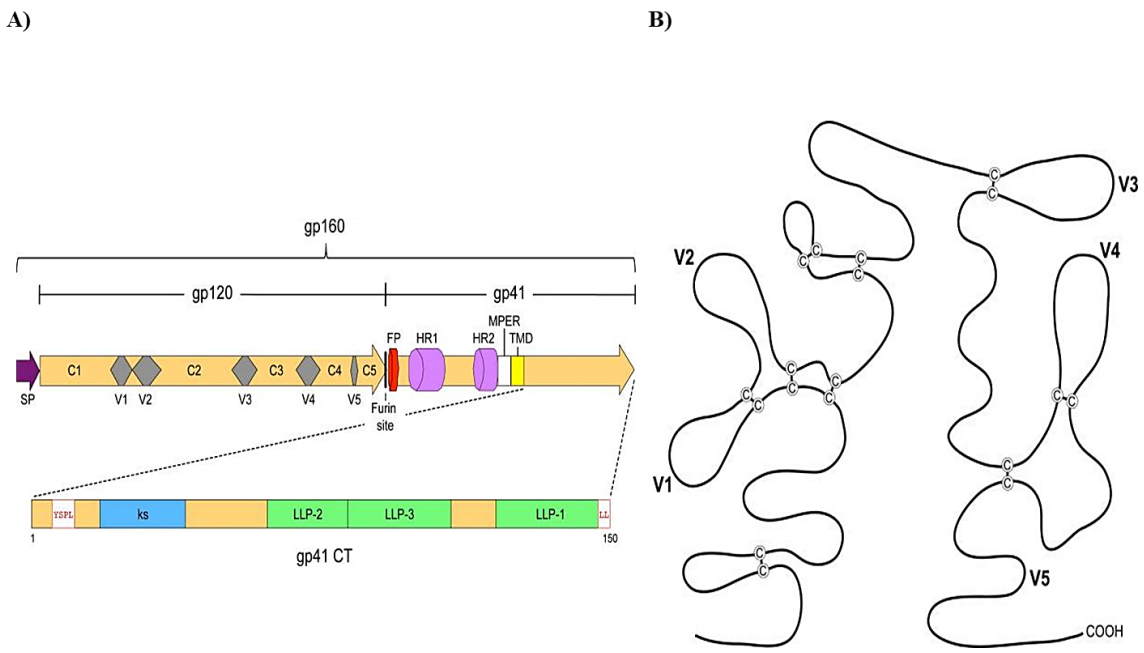


Figure 1.8. Schematic of Envelope Structure. A) Diagram of Envelope showing gp160 comprising gp120 and gp41 subunits. The variable regions are indicated by grey diamonds, the fusion peptide is indicated in red, HR1 and HR2 are in purple and the transmembrane domain (TMD) is in yellow. Gp41 consists of the internalization signal (YSPL); the Kennedy sequence (ks); the conserved amphipathic α-helices LLP-1, LLP-2, and LLP-3; and a C-terminal dileucine motif (LL). B) Gp120 linear structure showing the five variable loops with disulphide bonds indicated in grey with the constant regions between the variable loops (Checkley et al., 2011).

.

1.8.1.1 Entry of CD4+ cells

Entry of CD4+ cells can be summarized in three steps: 1) gp120 - CD4 binding, 2) gp120–chemokine receptor binding and 3) membrane fusion facilitated by gp41 (Figure 1.9). Gp120 binding to the CD4 receptor stimulates conformational changes and rearrangement in gp120, that allows exposure of the coreceptor binding sites necessary for gp120 binding to CCR5 or

CXCR4 (Wilen et al., 2012b). Initial gp120 conformational changes involve the shifting of the V1/V2 stem to expose the V3 loop coreceptor binding site, followed by the formation of the bridging sheet, the second coreceptor binding site (Koning et al., 2002; Rizzuto, 1998). Binding to the coreceptor involves two interactions: firstly, the crown of the V3 binds to the extracellular loop-2 (ECL2) of the CCR5 coreceptor followed by binding of the four-stranded bridging sheet and the base of V3 with the tyrosine sulfate of the CCR5 N- terminal domain (Huang et al., 2007). Then, each gp120 monomer rotates and partially reveals the gp41 stalk. Furthermore, Env-CD4 interaction is also associated with changes in the CD4 receptor which bring the virus and target cell membranes closer (Kwong et al., 1998; Liu et al., 2008; Wilen et al., 2012a).

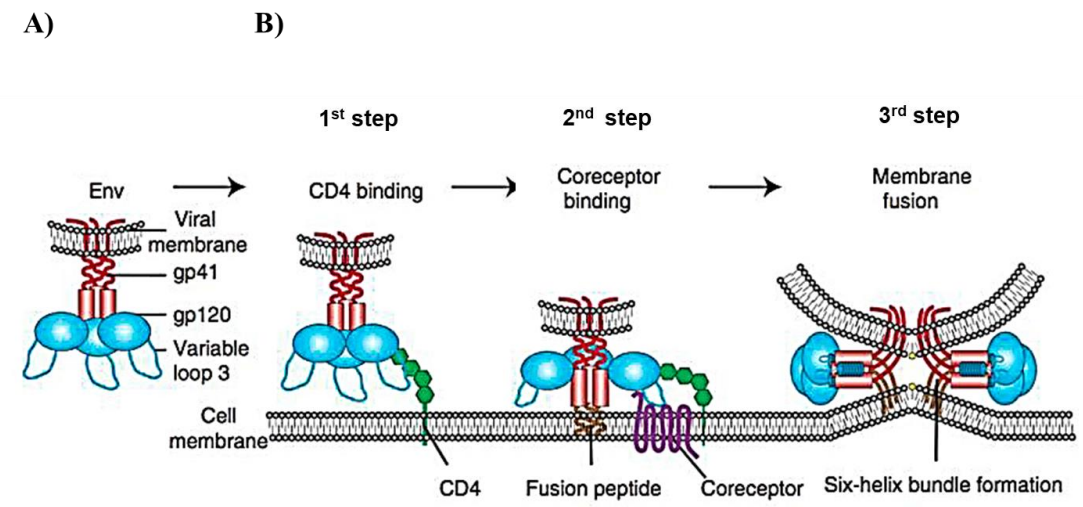


Figure 1.9. HIV entry process. A) HIV-1 Envelope inserts into the membrane via gp41 with the cytoplasmic tail facing into the cell. B) The virus entry mechanism is summarized by three steps: 1) CD4 binding resulting in conformational change, 2) V3 loop binding to coreceptor resulting in insertion of fusion peptide into the host cell membrane, and 3) Formation of the six-helix bundle drawing the viral and cell membranes close together for fusion. Figure adapted from Wilen et al. (2012b).

After gp120 binding to the coreceptor, additional conformational changes in gp120 and gp41 trigger the exposure of the FP of gp41 that inserts into the cell membrane, destabilizing it and resulting in the formation of a fusion pore that culminates in the fusion of host and viral

membranes and the delivery of the viral core (capsid) into the cytoplasm (Checkley et al., 2011; Liu et al., 2008; Yue et al., 2009). The FP of each gp41 then folds and brings the N-terminal HR-1 together to form a triple-stranded coiled coil structure. Subsequently, the C-terminal HR-2 from each gp41 subunit folds back and packs into a groove of the triple-stranded coiled coil. The formation of the 6HB stabilizes and stimulates the fusion pores to expand (Markosyan et al., 2003, 2009) and juxtapose the viral and host cell membranes leading to fusion (Figure 1.9). Both cell-cell fusion and virus-cell fusion can occur *in vivo* and both processes are mediated by Env (Freed and Martin, 1995a).

Env coreceptor binding sites have different affinities for CCR5 and CXCR4 which determine coreceptor tropism. The V3 region, with its positively charged residues is the main region responsible for determining viral tropism. Viruses that utilize CCR5 for entry are called R5 viruses and have a net V3 charge of +3 to +5. These viruses are usually responsible for virus transmission, and are found during the early stage of HIV infection whereas subtype B viruses that use CXCR4 i.e. X4 viruses with a net V3 charge of +7 to +10, emerge late during infection. Viruses that are able to bind to both CCR5 and CXCR4 for entry are designated as dual tropic (Hartley et al., 2005; Koning et al., 2002; Staropoli et al., 2000). The emergence of X4 and R5X4 tropic viruses later in infection is associated with a decline in CD4+ T cells and increased disease progression (Connor et al., 1997; Schramm et al., 2000). There is an important distinction between coreceptor usage and cellular tropism (Gorry and Ancuta, 2011). Viruses that infect macrophages are called macrophage-tropic viruses (M-tropic), while viruses that are able to infect CD4+ T cells are called T-lymphocyte tropic (T-cell tropic) even though both M-tropic and T-tropic viruses utilise CCR5. TFs are not M-tropic (Nofemela, 2013; Ping et al., 2013) although the ability to infect macrophages over the course of infection increases. The distinction between M-tropism and T-tropism is the ability of M-tropic variants to bind to low levels of CD4 on macrophages (Gorry et al., 2004; Gray et al., 2005) whereas T-tropic virus requires high levels of CD4 associated with T lymphocytes (Gorry et al., 2014; Ping et al., 2013). Therefore, viruses that are transmitted require high levels of CD4 for infection.

1.8.1.2 Envelope role in HIV-1 escape immune responses

Env is heavily glycosylated (Figure 1.10) carrying about 24 to 28 PNGs defined by the sequence motif: NXS or NXT where X is any amino acid other than proline (Lavine et al., 2012). Some PNGs are highly conserved in most subtypes, such as N241 within C2 on gp120, while others are highly conserved in some subtypes and not in others. For example, N339 within C3 of gp120 is highly conserved in HIV-1 subtype BC and B with 100-83%, respectively, but not in subtype AE (Wang et al., 2013). The high density of N-glycans plays many significant roles in Env function. N-glycosylation influences Env folding and processing and thus Env incorporation, virus entry and infectivity (Wang et al., 2013), receptor and coreceptor binding (Pollakis et al., 2001) and provides a "glycan shield" that protects against host immune responses (Moore et al., 2012; Townsley et al., 2016). More recently, it has been found that these N glycans are targeted by broadly neutralising antibodies (bnAbs) (Doores, 2015).

Env is under constant selective pressure by the host immune system and the addition and deletion of PNGs is a common mechanism to escape neutralisation (Wei et al., 2003) while contributing to the high diversity between HIV-1 variants. More recently, a study showed a shift in PNG position from 334 to 332 after 6 months of infection to escape early strain-specific antibodies (Moore et al., 2012), confirming the important role of N-glycosylation in evasion of immune responses. The N-glycan at position 332 was mapped to the gp120 outer domain within the epitope for a bnAb, 2G12, that is found within the glycan patch (Figure 1.11) (Lavine et al., 2012; Murin et al., 2014).

A fraction of HIV-1 infected individuals are able to generate bnAbs against diverse primary isolates (Pejchal et al., 2011). These bnAbs take years to evolve a high level of neutralisation breadth (Burton and Mascola, 2015). bnAbs target different regions in HIV-1 Env such as VRC01 that recognises the CD4-binding site of gp120 and 4E10 and 10E8 bnAbs that target the MPER on gp41 (Figure 1.11) (Doores, 2015). Interestingly, 2G12, interacts with multiple N-linked glycans, including N295, N332, N339 and N392 within the highly immunogenic "mannose patch" (Murin et al., 2014), suggesting that N-glycans contribute to nAb epitopes. In fact, the appearance of the PNG at site 332 6 months after infection not only allowed

escape from nAb but itself became the target for highly potent bnAb such as mAbs PGT121–PGT123, PGT125–PGT128, PGT130, PGT131 and PGT135–PGT137 (Moore et al., 2012). Therefore, although Env N-glycans protect the virus from neutralising antibodies, they are also the target for highly potent BCN antibodies (Binley et al., 2010).

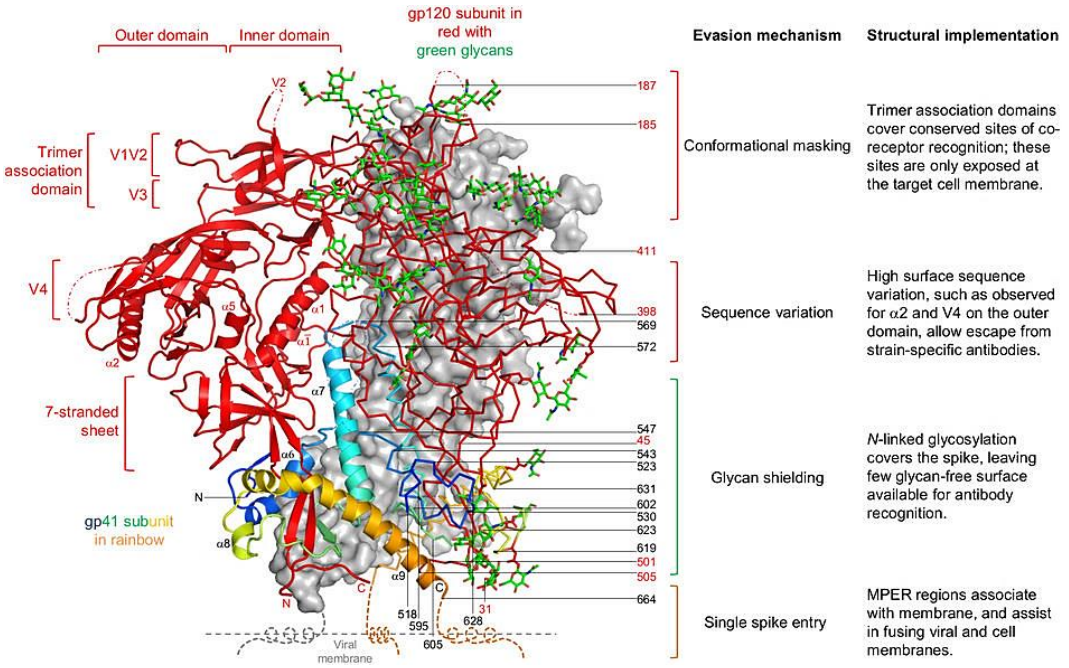


Figure 1.10. HIV-1 Env trimer. The three heterodimers are distinguished by stick, ribbon and grey surface representation on the left and a table of evasion mechanisms are indicated on the right. Gp120 is in red and gp41 is multi-coloured (from a blue N terminus to an orange C terminus), and crystallographically defined glycans are in green". The MPER of each heterodimer is shown by broken lines juxtaposed along the viral membrane. From "Structure and immune recognition of trimeric pre-fusion HIV-1 Env" (Pancera et al., 2014).

Deletion of N332 (and N386) resulted in a significant disruption of entry of YU-2 virus subtype B but not of JRFL of the same subtype (Lavine et al., 2012). However, the same PNGs were not essential for virus entry of clade BC virus (Wang et al., 2013), suggesting that the effect of PNGs on virus entry/infectivity can vary dramatically between subtypes and even between isolates of the same subtype. The loss of PNGs at V1/V2 and C1/C2 of gp120 resulted in significant loss of virus infectivity (Wang et al., 2013) possibly because PNGs maintain the steric conformation of HIV-1 Env (Huang et al., 2012). However, other

PNGs were reported to enhance viral activity such as N408 and N411 located in V4 of gp120 (Wang et al., 2013). Interestingly, deletion of some PNGs, such as N355 in C3 of gp120 and N611 in gp41, had no significant effect on viral entry despite causing a significant reduction in gp120 incorporation into virions (Wang et al., 2013). Overall, PNGs not only mask epitopes as a mechanism to escape the immune system, they also play a crucial role in Env entry and viral fitness.

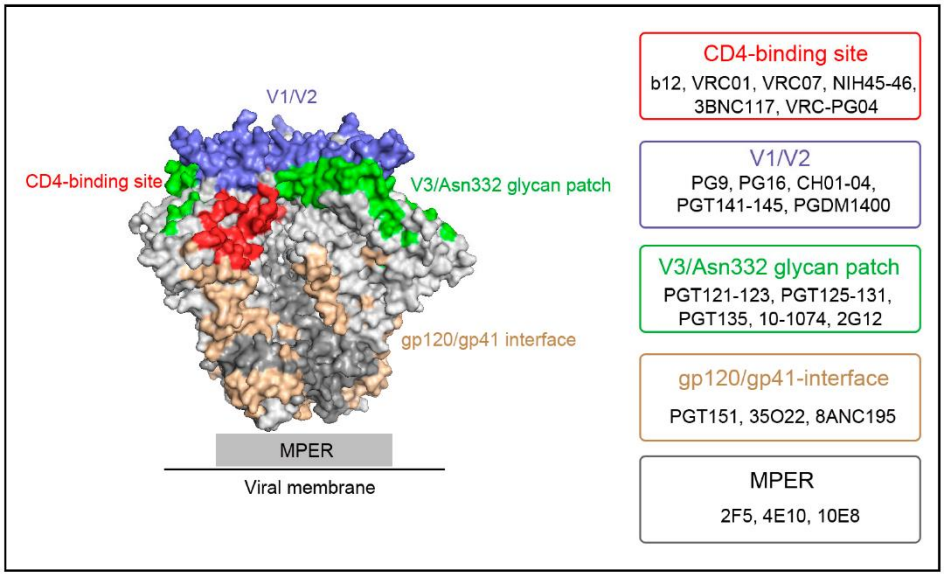


Figure 1.11. Location of broadly neutralizing antibody (bNAbs) epitopes on the Envelope trimer. Antibodies are listed on the right that recognise HIV-1 Env epitopes highlighted in different colours on the left [CD4bs (red), V1/V2 (blue), V3/Asn332 glycan patch (green), gp120/gp41-interface (brown), MPER (denoted as a grey rectangle)]. The surface area of gp120 and gp41 is highlighted in light grey and in dark grey, respectively. From Zhang et al, (2016). (Zhang et al., 2016).

1.8.1.3 Env expression and incorporation into viral particles

During virus assembly, Env trimers are incorporated into viral particles (Lambele´ et al., 2007), an essential step in the formation of infectious virions (Affranchino and González, 2014; Freed and Martin, 1996).

Env glycoproteins are synthesized, processed and glycosylated in the lumen of the RER. After oligomerisation of gp160, the trimers traffic through the Golgi complex where N-

glycans are further modified and proteolytic cleavage of gp160 into gp120 and gp41 occurs *en route* to the plasma membrane (PM) (Freed et al., 1989; McCune et al., 1988) or *trans*-Golgi-derived (TGN) secretory vesicles (Moulard and Decroly, 2000). Cleavage of gp160 is an absolute requirement for Env fusogenicity. Gp120 and gp41 remain non-covalently connected and once at the PM, the heterotrimers are incorporated into viral particles. Studies have consistently shown that the number of heterotrimers incorporated per virus is very low: ~7-14 (Checkley et al., 2011; Murakami, 2008) and ~ 10–14 spikes/virion (Didigu and Doms, 2012; Staropoli et al., 2000; Wyatt, 1998).

Env incorporation requires interaction between gp41 CT and the N-terminus of p55Gag precursor- MA (Affranchino and González, 2014; Freed and Martin, 1995b; Murakami, 2008; Murakami and Freed, 2000a; Postler and Desrosiers, 2013; da Silva et al., 2016; Wyma et al., 2000). For instance, a single mutation in MA restored Env incorporation that was lost due to a small deletion in the LLP2/LLP3 region of gp41-CT (Murakami and Freed, 2000a). Conversely, a mutation in MA that disrupted Env incorporation was reversed by truncating a small region of gp41 (Freed and Martin, 1995b; Mammano et al., 1995). Therefore, gp41 CT is essential for Env incorporation (Affranchino and González, 2014; Da Silva et al., 2013; Freed and Martin, 1995b; Murakami, 2008; Murakami and Freed, 2000b; Postler and Desrosiers, 2013; Wyma et al., 2000).

Checkly *et al*. (2011) summarised several proposed Env incorporation mechanisms: 1) The "passive" or "random" incorporation model (Figure 1.12A) suggests that Env trafficking to the PM takes place independently of Gag and Env is incorporated into the virion as a result of Env expression and virus assembly; 2) The "direct Gag–Env interaction" model (Figure 1.12B) where a direct interaction between MA and gp41-CT is required for Env incorporation; 3) The "Gag–Env co-targeting" model (Figure 1.12C) which describes an indirect interaction between Env and Gag in which cellular host factors (lipid rafts) play a role in Env and Gag enrichment at the PM enabling virus assembly; and 4) The "indirect Gag–Env interaction" model (Figure 1.12D) when other cellular proteins (adaptor proteins) link Env and Gag and stimulate Env incorporation (Checkly et al. 2011).

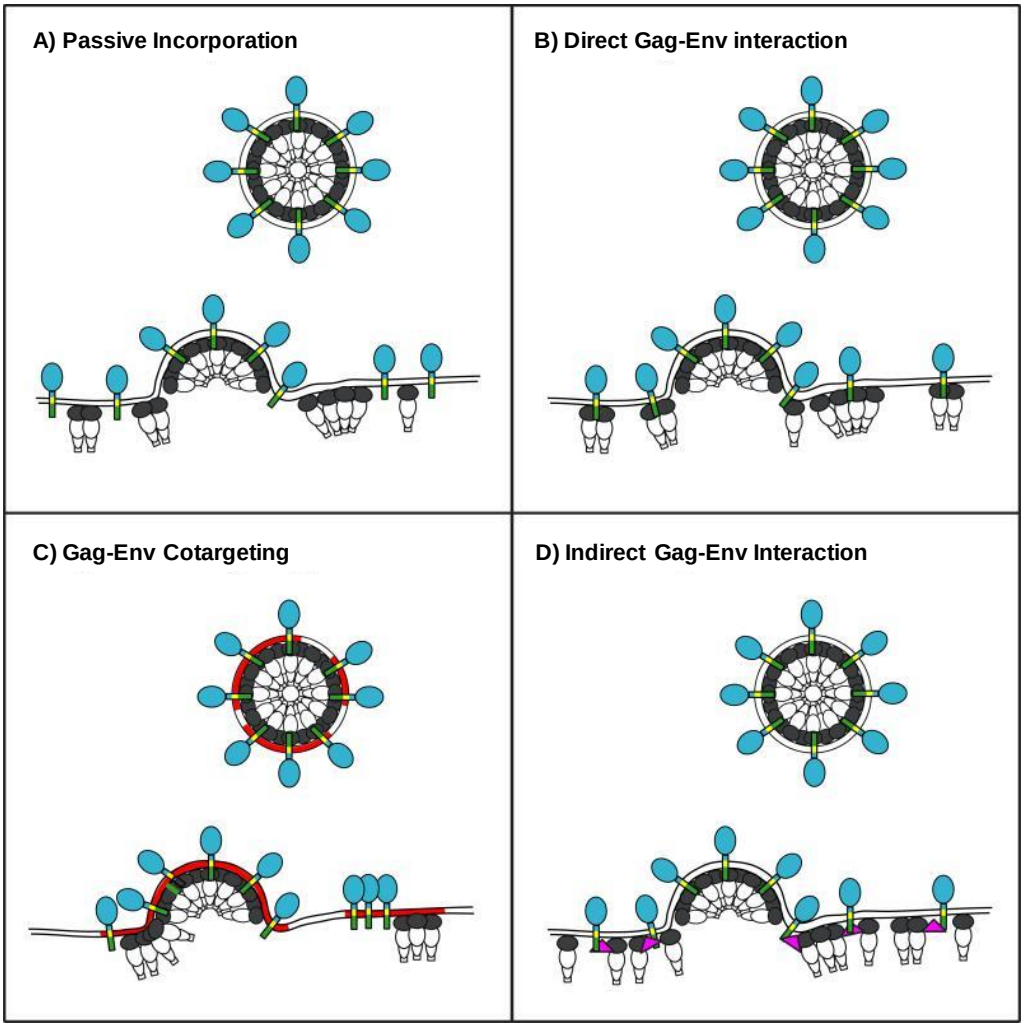


Figure 1.12. Env incorporation models explains. A) Passive model where interaction between Gag and Env is not required, B) direct interaction between Gag-Env is required; C) co-targeting of Gag and Env for recruitment to the PM via lipid drafts; D) Gag-Env interaction requires an adaptor protein to link Env and Gag together. Env complex spikes, Gag, lipid drafts, and the adaptor protein are in indicated by blue, grey, red and pink color, respectively. Figure was taken from Checkley et al. (2011).

1.9 Viral fitness

Viral fitness is defined as the ability of the virus to adapt and replicate in a given host environment (Domingo and Holland, 1997). Recently it was shown that participants that became infected with TFs with high replication capacity, progressed more rapidly to disease and carried a larger early latent viral reservoir (Claiborne et al., 2015). As transmitted variants tend to be more pathogenic than previously believed, it was suggested that viral fitness could be used as a clinical marker for the start of ART given its association with disease progression (Arnott et al., 2010).

1.9.1 Env as an HIV fitness determinant

Although Gag has been shown to play an important role in HIV-1 disease progression (Prince et al., 2012), there is evidence as to the importance of Env in cell entry and tropism (Hoffman, and Doms, 1999; Pastore et al., 2006; Pollakis et al., 2001; Wilen et al., 2012a), transmission (Hsu et al., 2003), and virus fitness and disease progression (Lobritz et al., 2007; Quiñones-Mateu et al., 2000). Moreover, Env plays a key role in escape from immune responses and virus survival due to its high glycosylation (Moore et al., 2012; Wang et al., 2013; Wei et al., 2003). Rangel et al. (2003) determined a significant correlation ($p = 0.002$) between the fitness of wild-type HIV-1 isolates and the entry efficiency of their corresponding Env using an *ex vivo* competition assay. He also showed in a time course competition assay that virus entry had a higher impact on the ability of one virus to outcompete and dominate another compared to other steps in the HIV life cycle (Rangel et al., 2003). In another competitive study, a primary HIV isolate had similar fitness to its *env*/NL4-3 chimeric counterpart (Marozsan et al., 2005). A more recent study, using a similar approach, compared the fitness of an isolate from a multi-drug resistant (MDR) participant to its corresponding *env*/NL4-3 chimera and a control construct comprising the isolate's backbone genome carrying Env from pNL4-3. The result showed that both the isolate and its Env had higher fitness than its backbone genome only, suggesting that Env is a major determinant in viral fitness (Mohri et al., 2015). Similarly, variants resistant to MVC carried a mutation within the CD4 binding site that rendered them better at replication than MVC sensitive clones (Ratcliff et al., 2013). As Env plays a very important role in immune evasion, variants resistant to neutralisation by

VRC01 were tested for changes in host cell entry and replication fitness. CD4-mediated viral entry and replicative capacity was diminished in the presence of Env mutations that conferred resistance to VRC01 (Lynch et al., 2015). Therefore, the relationship between Env fitness and changes in viral replication could be due to variation in the interaction between Env with CD4 and/or CCR5 (Marozsan et al., 2005) as well as other receptors. Env binds DCs via mannose-binding lectins such as dendritic cell–specific intercellular adhesion molecular 3-grabbing non-integrin (DC-SIGN) and interacts with α4β7 integrin on CD4 T cells (Arthos et al., 2008; Baribaud et al., 2001; Mondor et al., 1998). Binding to these receptors facilitates the transfer of infectious virus from the genital tract to the draining lymph nodes, facilitating productive, clinical infection. Binding of Env to DC-SIGN and α4β7 has been suggested to play an important role in HIV-1 transmission (Arthos et al., 2008; Geijtenbeek et al., 2002). Therefore, changes within the structure of Env could influence these interactions and thus its transmission fitness.

1.9.2 Env fitness and disease progression

HIV-1 Env fitness was suggested to play an important role in disease progression (Connor and Ho, 1994). A study comparing relative fitness of viral isolates from LTNP and rapid progressors, demonstrated an association between disease progression and viral fitness (Quiñones-Mateu et al., 2000). Similarly, Lassen et al. (2009) showed that Env clones derived from elite suppressors displayed low Env entry efficiency compared to Env derived from chronic progressors (Lassen et al., 2009). Troyer et al., (2005), showed a strong correlation between *ex vivo* fitness of HIV-1 isolates and both viral diversity and disease progression (Troyer et al., 2005) whereas Gordan et al. (2016) reported high *ex vivo* entry efficiency of an *env* pseudotyped virus potentially associated with rapid disease progression (Gordon et al., 2016). Recently, it was suggested that the apoptosis-inducing potential of Env might be responsible for the depletion of CD4+ T cells in HIV-1 infection, directly linking Env with disease progression (Joshi et al., 2016). Finally, a MDR isolate, shown to have high Env entry efficiency, was associated with rapid disease progression (Mohri et al., 2015). Therefore, escape from immune responses and antiretroviral therapy could select for mutations in Env that enhance entry efficiency which, in turn, drives viral fitness and disease

progression. Overall, this review has highlighted the importance of Env, diversity, expression, processing and function in viral fitness and disease progression.

1.10 Research Aim and objectives

Use dual infection as a model to understand the role of Envelope entry efficiency in the *in vivo* outgrowth of variants, the replicative capacity of infectious molecular clones and disease progression.

1.10.1 Rational

Dual infected individuals provide a unique model system for the identification of determinants of viral fitness as two distinct viruses co-evolve under the same environmental pressures culminating in the outgrowth of recombinant variants most likely due to gain of fitness mutations. *The overall hypothesis tested in this study will be that rapid viral evolution in dual-infected individuals results in the rapid generation of recombinants with high replication fitness that drive disease progression*. Therefore, recombinants with high replication fitness that emerge early in infection could carry regions/sites on Env that are essential for viral propagation. These regions can be used to guide drug and immunogen design. This study aims to identify and characterise determinants of subtype C HIV-1 Env fitness in individuals co-infected with two phylogenetically distinct viruses to evaluate their impact on viral replication in PBMCs. This is a longitudinal study of the genotypic and phenotypic evolution of Env isolated from four HIV-1 subtype C dual-infected South African women differing in disease progression.

1.10.2 Objectives

I. Determine the association between *in vitro* Env entry efficiency and *in vivo* outgrowth of variants infecting dual infected individuals and identify potential functional determinants that could be targeted in drug and vaccine design

II. Understand how the function of Env fitness determinants could impact the *in vivo* outgrowth of variants isolated from dual infected individuals
III. Determine whether Env entry efficiency impacts the replicative capacity of chimeric infectious molecular clones in PBMCs.
IV. Describe whether there is an association between Env entry efficiency, IMC replicative fitness and disease progression.

CHAPTER 2.

The Role of HIV-1 Subtype C Envelope Entry Efficiency In *In Vivo* Outgrowth of Variants Infecting Dual-Infected Individuals

2.1 Introduction

In dual infections the presence of two distinct viruses within the same host and under the same selection pressures could lead to "*in vivo"* competition, increasing the probability of rapid evolution, high diversity of circulating variants and potentially reducing the efficacy of vaccines (Sanborn et al., 2015). Dual infections are also likely to lead to the emergence of recombinant variants with advantageous mutations, inherited from the two initial viruses, that enhance viral replication fitness (van der Kuyl and Cornelissen, 2007; Peeters, 2000).

Studies have shown that multiple variant infections were significantly associated with increased viral load, CD4+ T cell decline and thus rapid disease progression (Cornelissen et al., 2012; Gottlieb et al., 2004; Sagar et al., 2003). A recent study showed that infection with multiple variants correlated with a significantly higher viral load after the first year of infection compared to patients infected with a single variant (Janes et al., 2015). It was also suggested that the rapid disease progression of a dual infected individual to AIDS defining illness was due to the emergence of a recombinant variant within the first year of infection with enhanced Envelop (Env) entry efficiency compared to its parental variants. The authors suggested that Env might play an essential role in the outgrowth of variants with enhanced viral fitness that could potentially increase disease progression (Gordon et al., 2016).

Env entry efficiency has been reported to be associated with disease progression (Campbell et al., 2003; Gordon et al., 2016; Quiñones-Mateu et al., 2000; Troyer et al., 2005). Recently, a study showed that Env entry efficiency of transmitted founder (TF) variants impacted viral loads during acute infection (Nofemela et al., 2015). HIV Env mediates the first step in virus

entry of host cells by interacting with cell-surface receptors e.g. CD4, CCR5 and CXCR4 (Wilen et al., 2012a). Env is subjected to extensive selective pressure, including host immune response [humoral and cytotoxic T-lymphocyte (CTL)] (Troyer et al., 2005), levels of co-receptor expression (Gorry et al., 2004; Meijerink et al., 2014) and availability of permissive host cells (Marozsan et al., 2005). Reports support the hypothesis that Env plays a major role in the competitive and adaptive ability of the virus (Ball et al., 2003; Rangel et al., 2003). Together, the data suggest that Env entry efficiency (Env fitness) might play an important role in the outgrowth and thus the pathogenicity of variants infecting dual infected individuals.

Viral fitness, the ability to survive within a given environment includes the ability of variants to evade the host immune system and resist antiretroviral therapy (ART). Therefore, viruses able to evade host immunity or carrying drug resistant mutations can result in loss of viremic control due to treatment failure or compromised immune system potentially leading to rapid disease progression and possibly death of the patient (van der Kuyl and Cornelissen, 2007). One solution is to restrict variants from introducing fitness-enhancing mutations by designing vaccines and drugs that target sites/regions essential for HIV survival so that any amino acid change either leads to loss of fitness or viral clearance (Arenas et al., 2016; Sather et al., 2012). However, to accomplish this strategy, we need to identify HIV-1 viral fitness determinants.

Dual-infected individuals provide a unique model system to characterise the evolution of Env function over the course of infection as two phylogenetically distinct viruses co-evolve under the same environmental pressures. Therefore, we aimed to track Env entry efficiency of variants isolated from dual infected individuals and evaluate its impact on virus competition and disease progression and thereby identify Env fitness determinants.

2.2 Research Aim and Objectives

Aim

The overall aim of this study was to investigate whether the entry efficiency of HIV-1 subtype C Env influenced the outgrowth of variants infecting dual infected individuals and to identify potential functional determinants.

Objectives

Objective 1: Determine the *in vivo* frequency of viral populations infecting dual infected individuals using sequence analysis of longitudinal samples

Objective 2: Determine *in vitro* Env entry efficiency of variants over time

Objective 3: Identify regions/sites within Env involved in replication fitness of viruses infecting dual infected individuals.

2.3 Materials and Methods

2.3.1 Study cohort and samples used in this analysis

This study utilized stored PCR products generated as part of the CAPRISA 002 Acute Infection study, Durban, South Africa (van Loggerenberg et al., 2008). HIV-positive women were enrolled within 3 months of infection and plasma samples were collected at enrolment (Phase 2_week/month post-infection_visit), then weekly for three weeks, fortnightly until 3 months, monthly until a year and quarterly thereafter. Pre-enrolment samples were identified by testing samples prior to enrolment for viral RNA (Phase 1, week/month post-infection_visit) (van Loggerenberg et al., 2008). In this study, we designated the first visit from which we obtained samples as 0 months irrespective if it was at enrolment or pre-enrolment sample.

Out of 39 individuals screened, only four CAPRISA 002 study participants were identified as dual infected (not superinfected): CAP37, CAP84, CAP137 and CAP267 (Woodman et al., 2011). One participant, CAP84, was classified as a typical progressor as the CD4+ T cell count remained stable; whereas CAP37, CAP137 and CAP267 were classified as rapid progressors as their CD4+ cell counts dropped to below 350 cells/µl within the first two years of infection (Mlisana et al., 2014). Viral loads were measured using a COBAS AMPLICOR^{TM} HIV-1 Monitor Test v1.5 (Roche Diagnostics, Rotkreuz, Switzerland) and CD4+ T cells counts were assessed using a FACS Calibur flow cytometer (BD Biosciences, San Jose, USA (Mlisana et al., 2014).

The SGA-derived first round, full-length *env* PCR products were obtained from Professor Williamson (Institute of Infectious Diseases and Molecular Medicine, University of Cape Town) (Abrahams et al., 2013). Briefly, viral RNA extracted from longitudinal plasma samples (0, 3, 6, and 12 months post infection (mpi) of CAP37, CAP84, CAP137 and CAP267, was reverse transcribed and cDNA was diluted so that PCR products were generated from a single template. Amplicons were directly sequenced and stored at –80°C for further analysis.

2.3.2 Analysis of SGA-derived envelope sequences

In the current study, a total of 295 full-length *env* sequences from all participants (with an average of 20 *env* sequences per time point) across time points were analysed using BioEdit version 7.0. Classification of variants was performed using the DNA pairwise distance analyses MEGA-5 (Tamura et al., 2011) in conjunction with recombination analysis using Highlighter plot (http://www.hiv.lanl.gov) (Kraft et al., 2012) and Recombinant Identification Program (RIP) (http://www.hiv.lanl.gov/content/sequence/RIP).

2.3.2.1 Identification of parent variants transmitted to dual infected individuals

Identifying the initial parent viruses transmitted to dual infected individuals was challenging due to the potential transmission of multiple variants from donors, early recombination between transmitted variants, high genetic diversity and limited detection of variants at low frequencies. To reduce the complexity of the number of different viral populations, the dominant viral population at enrolment (0 mpi) was designated as population A. Neighbour-joining trees (MEGA-5) were then constructed to identify variants that clustered separately to viral population A and the virus with the greatest DNA maximum distance was classified as the second 'parent', virus B. These two "parent" viruses were used as master sequences in Highlighter plots and all other viruses at 0, 3, 6 or 12 mpi were classified as either A, B or AB depending on their position on the Neighbour-joining tree, their DNA distance from each "parent" and their similarity compared to the master sequences on the matched Highlighter plots (http://www.hiv.lanl.gov), using the "match" setting to visualize the intra-participant diversity. In one or two variants, findings from Maximum DNA distance analysis, Highlighter plot, RIP and Neighbour-joining trees did not support one another and these sequences were classified according to RIP analysis only.

All recombinants were classified initially as AB even if there were more than one recombinant population. As the aim of the study was to track fit variants able to outcompete other viruses, we assumed that variants at low frequency were not driving viral loads in these participants and that undetectable viruses were not playing an essential role.

2.3.2.2 Frequency of variants infecting dual infected individuals over the course of infection

To track variants over the course of infection, viruses at all time points were compared to parents A and B using neighbour joining trees of sub-genomic regions, maximum DNA distance and Highlighter plots. Outgrowth of A, B or AB variants was determined by measuring the frequency of each viral population relative to the total number of sequences per time point (%).

After measuring outgrowth of AB variants, we looked at the recombinant populations in more detail to compare relative frequency and fitness of sub-groups. The classification of variants into subgroups of AB recombinants was confirmed by Highlighter plots (http://www.hiv.lanl.gov), RIP (http://www.hiv.lanl.gov/content/sequence/RIP) and Neighbour-joining trees.

2.3.3 Cloning functional Env SGA

2.3.3.1 Env amplicons selected for cloning

In order to investigate *in vitro* Env entry efficiency, SGA-derived *env* amplicons from each participant were cloned into mammalian vectors. Selection of *env* amplicons was based on 1) representing A, B and AB variants at each time point and 2) cloning representatives of populations that changed frequency over time. Where there were two recombinant sub-populations, representatives were cloned from both. For CAP37, eight functional Env clones representing viruses A, B and AB were generated: three clones from enrolment, two clones from 6 mpi and three clones from 12 mpi representing two AB populations. For CAP84, four functional Env clones were generated: one representing virus A at 0 mpi (84c1), one representing virus B at 3 mpi (84c2), and two representing virus A and AB at 12 mpi, (84c3, 84c4). Ten Env clones were generated representing variants infecting CAP137 over time. Two clones represent virus A at 0 mpi, one represents virus B at ~2 mpi, two representing A and B viruses at 3 mpi, three representing A, B and AB at 6 mpi, and two representing AB sub-population at 12 mpi. Six of the SGA-derived *en*v amplicons sequences were cloned for CAP267 and represent A and B viral populations at 0 mpi, 3 mpi and 12 mpi.

2.3.3.2 PCR Amplification of env

Representative SGA *env* amplicons were amplified using nested PCR performed by an XP Thermal cycler (Bioer technology). The *env* specific primers Env-1A-RX and Env-1M or Env-N (Appendix A) were used to amplify *rev-vpu-en*v expression cassettes as previously described (Kraus et al., 2010). PCR reactions were carried out according to the Phusion Hot Start Polymerase kit instructions (Thermo Scientific, USA) with some modifications described in Appendix A. The size of the PCR products was confirmed using 0.8 % agarose gel electrophoresis. The DNA band of approximately 3 Kb was excised and extracted from the agarose gel using the Wizard® SV Gel and PCR Clean-Up System (Promega, USA) following the manufacturer's instructions.

2.3.3.3 'A' tailing and Ligation Reaction

In order to clone the PCR products into a mammalian expression vector, pTargeT™, a poly-adenosine (poly-A) tail was added to both the 5' and 3' ends of the cleaned-up *env* PCR products using SuperTherm *Taq* polymerase enzyme (JMR- Holdings, UK). The reaction was carried out at 72 °C for 10 min and the reagents used for the poly-A tailing reaction are listed in (Appendix A). The poly-A-tailed PCR products were cleaned up using the Wizard SV PCR Gel clean-up kit (Promega, USA), and ligated with pTargeT™ vector following the manufacturer instructions (Promega, USA). The ligation reaction was performed overnight at 4 °C using T4 ligase (Promega, USA) at a molar ratio of 1:3 (vector: insert).

2.3.3.4 Transformation

The competent *E.coli* JM109 bacterial cells (Promega, USA) were used for transformation as follows: the bacterial cells were allowed to thaw on ice for 10 min and the ligation reaction was added to 50 µl competent cells at a 1:10 ratio of plasmid:cells. The mixture was incubated at 4 °C for 20 min, and then subjected to heat shock for 45 sec at 42 °C and then returned to ice for a further 2 min. The bacterial cells were allowed to recover at 37 °C for 1 hr in 950 µl Luria broth (LB) without Carbenocillin. The bacterial cells were concentrated by centrifugation at 5000 rpm for 7 min and the pellet was re-suspended in 50 µl of LB and

spread on Luria agar plates (1.5 % agar and 0.1 mg/ml Carbenocillin). The bacterial plates were incubated at 30 °C overnight.

Colony PCR was used to screen colonies for the presence of *env* in the correct orientation downstream the CMV promotor of the pTargeT™ vector (Promega, USA). A T7 primer complimentary to the T7 promoter sequence was used as forward primer and the Rev15 primer complimentary to *rev* sequence was used as the reverse primer. The Supertherm Taq polymerase enzyme (JMR- Holdings, UK) was used to amplify a product of 1 Kb using the PCR reaction mixture and cycling conditions listed in (Appendix A). Once the insert was verified, LB was inoculated with the bacterial colony and the plasmid was purified using the PureYield™ Plasmid Midiprep System (Promega, USA) as per manufacturer's instructions. Plasmids were sequenced at Stellenbosch Central Analytical Facility using an ABI 3000 genetic analyser (Applied Biosystems, USA). Sequence analysis to verify correct *env* sequence was performed using ChromasPro and compared to the *env* SGA-derived sequence.

2.3.4 Pseudovirus Entry Efficiency

In order to compare Env entry efficiency, pseudovirus (PSV) was generated using human embryonic kidney (HEK-293T) cells and entry efficiency was compared using the TZM-bl reporter cell line.

2.3.4.1 Cell culture

HEK-293T: Highly transfectable HEK-293T cells were (Graham et al., 1977) maintained in T75 flasks at 37 °C with 5 % CO_2 in Dulbecco's Minimal Essential Medium (DMEM) with high Glucose (4.5 g/L) obtained from Lonza (Switzerland) or Sigma (Germany). Medium was supplemented with 10 % Fetal Calf Serum (FCS) (BiochromGmBH), and 1 U/ml Penicillin and 1 µg/ml Streptomycin (Lonza, Switzerland).

TZM-bl cells: HeLa cell line that stably expresses CD4, CCR5 and CXCR4 and thus susceptible to HIV infection (Platt et al., 1998; Wei et al., 2002). It also expresses luciferase and β-galactosidase enzymes under the control of a Tat-inducible HIV LTR promoter. Cells were maintained as described above.

2.3.4.2 Generating pseudovirus and entry efficiency assay

PSV was generated by co-transfecting HEK-293T cells with 2.5 µg of *env* plasmid and 5 µg of pSG3ΔEnv HIV-1 backbone vector (1:1.5 molar ratio) (NIH AIDS Reagent Program, Division of AIDS, NIAID, NIH: from Drs. John C. Kappes and Xiaoyun Wu: pSG3Δenv) using 1 mg/ml Polyethylenimine (PEI). PSV supernatant was collected after 48 hr and stored at -80 °C in 20 % FCS prior to determining PSV titre using p24 ELISA (section 2.3.4.3). Using 96-well plates, (10^4 cells / 100 µl) of TZM-bl cells (AIDS Research and Reference Reagent Program, Division of AIDS, NIAID, NIH; from Dr. John C. Kappes, Dr. Xiaoyun Wu, and Tranzyme Inc.) were infected in triplicate with PSV normalized to 100 ng/ml p24. After 48 hrs, medium was replaced with Bright-Glo Luciferase Assay buffer (Promega, USA) and relative luciferase units (RLU) were measured using a GloMax-Multi Microplate Multimode Reader (Promega, USA). RLU readings of uninfected TZM-bl cells were considered as background signal and only RLU readings 2.5-fold higher than background was considered as positive infection. PSV generated by transfecting the HIV-1 subtype B pSG3ΔEnv backbone vector with pTargeT™ vector only (with no *env* cloned) was used as a negative control.

2.3.4.3 P24 antigen ELISA

A modified p24 sandwich ELISA (Aalto Bio reagents) and TROPIX® detection system (CDP-Star®, Applied Biosystems) were used to determine PSV titre as previously described by Moore *et al*. (1990) and Mckeating *et al*. (1991) and adapted by Dr A Trkola (McKEATING et al., 1991; Moore et al., 1990; Trkola et al., 1995).

High binding 96 well plates (Porvair Sciences, UK) were coated with a 100 µl/well of the D7320 sheep polyclonal anti-HIV-1 p24 antibody (Aalto Bio reagents, Dublin, Ireland) diluted in 0.1 M NaHC03 buffer, pH 8.5 to a final concentration of 2 µg/ml and incubated overnight at room temperature. Unbound antibody was removed by washing the plates three times with 1 X Tris-buffered saline (TBS). Plates were blocked with 5 % bovine serum albumin (PAA, Biocom Biotech), dissolved in TBS for 1-2 hrs at room temperature and thereafter stored in blocking agent at -20 °C until needed. PSV was inactivated by 5-fold dilution with 1.25 % Empigen detergent (Sigma-Aldrich, MO, USA) in 1 X TBS for one

hour at room temperature. Then serial dilutions of PSV were made using 1 % Empigen-TBS. A volume of 100 µl diluted samples was added to each well of pre-coated ELISA plates that were washed four times to remove the blocking reagent. The p24 standard curve was generated by 2-fold serial dilution of recombinant HIV-1 p24 antigen standard (Aalto Bio Reagents, Dublin, Ireland) starting at 32 ng/ml and ending at 0.5 ng/ml. Then, each p24 dilution was loaded in duplicate onto each ELISA plate (100 µl/well). Plates were incubated for two and half to three hours at room temperature to allow p24 antigen to bind to the antibody coated on the surface of the wells and thereafter washed three times with 1X TBS. The secondary anti-HIV-1 p24 conjugate (BC 1017-AP, Aalto Bio Reagents, Dublin, Ireland) was prepared by diluting the antibody 1: 64000 with 0.1 % Tween-20 (TBS-T) containing 20 % sheep serum (PAA, Biocom Biotech). A volume of 100 µl of the diluted antibody was added to each well in the plate. The secondary antibody was allowed to bind for one hour at room temperature followed by three washes with (1X TBS-T) and then two washes with 1X TROPIX buffer (Appendix C). For detection, 100 µl of CPD-Star® with Sapphire II enhancer substrate (Applied Biosystems, CA, USA) diluted 1:4 in 1 X TROPIX buffer was added to each well and incubated for 10 min at room temperature in the dark after which luminescence (measured in RLU) was read immediately using a Glomax luminometer (Promega, MA, USA).

2.3.5 Env determinants for viral fitness

In order to identify functional Env determinants of entry efficiency, chimeras and mutants were generated from selected participants.

2.3.5.1 Generating Chimeras

The *env* sequences generated as described in section 2.3.3 were aligned using BioEdit and the Gene cutter tool (http://www.hiv.lanl.gov) was used to identify different regions of Env (Appendix B).

CAP84 chimeras were generated by swapping the gp41 region between virus A clone 84c1 and virus AB clone 84c4 using restriction enzyme digestion (Figure 2.1). Theoretical constructs and restriction enzyme mapping were performed using Serial cloner 2.5 software.

The plasmid (5 µg) of the 84c1 and 84c4 were sequentially digested using *Sbf*I and *Xma*I restriction enzymes (Thermo Scientific, MA, USA) according to the manufacturer's recommendation. The required DNA bands were excised and gel purified using Wizard SV PCR Gel clean-up kit (Promega, USA) following the manufacturer's instructions. Ligation of the fragments was performed overnight by T4 ligase (Thermo Scientific, MA, USA) in 10 µl final reaction volume. Transformation was carried out using JM109 cells as described in section 2.3.3. Restriction enzyme digestion was used to screen colonies for mutagenesis and confirmed by sequencing.

CAP267 chimeras were generated by swapping gp120 between virus B at 0 mpi (267c2) and virus B at 12 mpi (267c6) using *Eco*N1 which cut twice in both clones before gp41 region (Figure 2.2). Colony PCR was used to confirm correct directional cloning after ligation followed by sequencing to ensure the correct constructs at Stellenbosch Central Analytical Facility using an ABI 3000 genetic analyser (Applied Biosystems, Foster City, CA, USA).

2.3.6 Site direct mutagenesis

Potential N-glycan sites (PNGs) were introduced in chimera 84c1/c4 (Figure 2.1) or removed from 84c4 and 137c10 (Figure 2.1, Figure 2.2) using a modified Stratagene Quik Change II PCR-based protocol (Agilent Technologies, CA, USA). Overlapping primers with silent restriction enzyme sites for screening were designed using WATCUT (http://watcut.uwaterloo.ca) and SDM-Assist software (Karnik et al., 2013). Briefly, mutated PCR-generated plasmids using the Phusion Hot Start Polymerase kit (Thermo Scientific) were digested with methyl-specific *Dpn*I- (Thermo Scientific/Fermentas, USA). JM109 competent cells (Promega, USA) were transformed with the un-methylated PCR products and screened using restriction enzyme digestion. Positive colonies were selected and plasmids were extracted using PureYield™ Plasmid Midiprep System (Promega, USA) as mentioned above and sequenced at Stellenbosch Central Analytical Facility using an ABI 3000 genetic analyser (Applied Biosystems, Foster City, CA, USA). All Env alignments are in Appendix B.

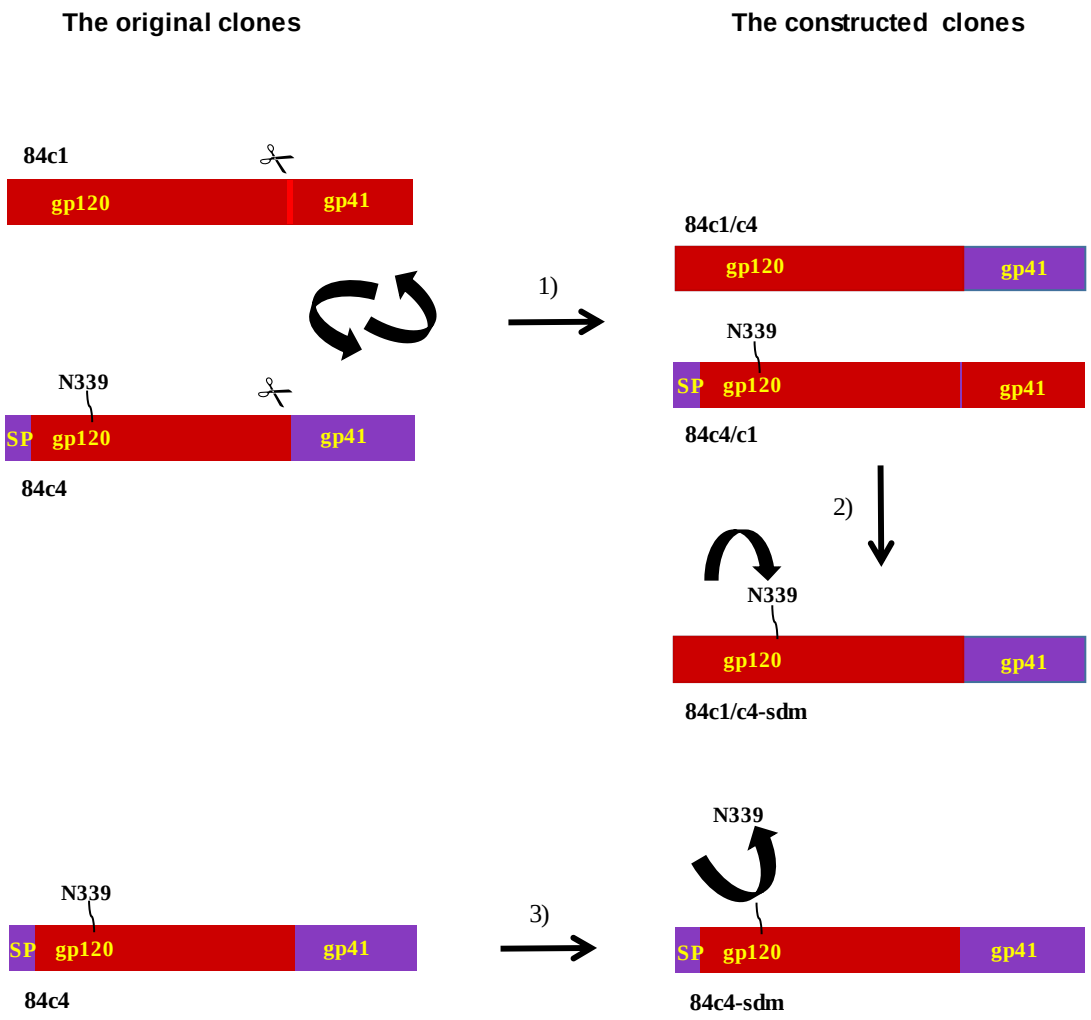


Figure 2.1. CAP84 chimeric Env clones and mutants. A diagram illustrating the construction of env chimeric clones and mutants generated from CAP84. 1) Two chimeras were constructed by swapping gp41 between 84c1 and 84c4 indicated by red and red with purple signalling peptide (SP) & gp41, respectively. 2) A potential N-glycan site was introduced at 339 using site direct mutagenesis (SDM) of gp120 of the chimera 84c1/c4 to generate a mutant 84c1/c4-sdm. 3) The potential N-glycan site at 339 presents in clone 84c4 was removed by SDM and the product was designated 84c4-sdm.

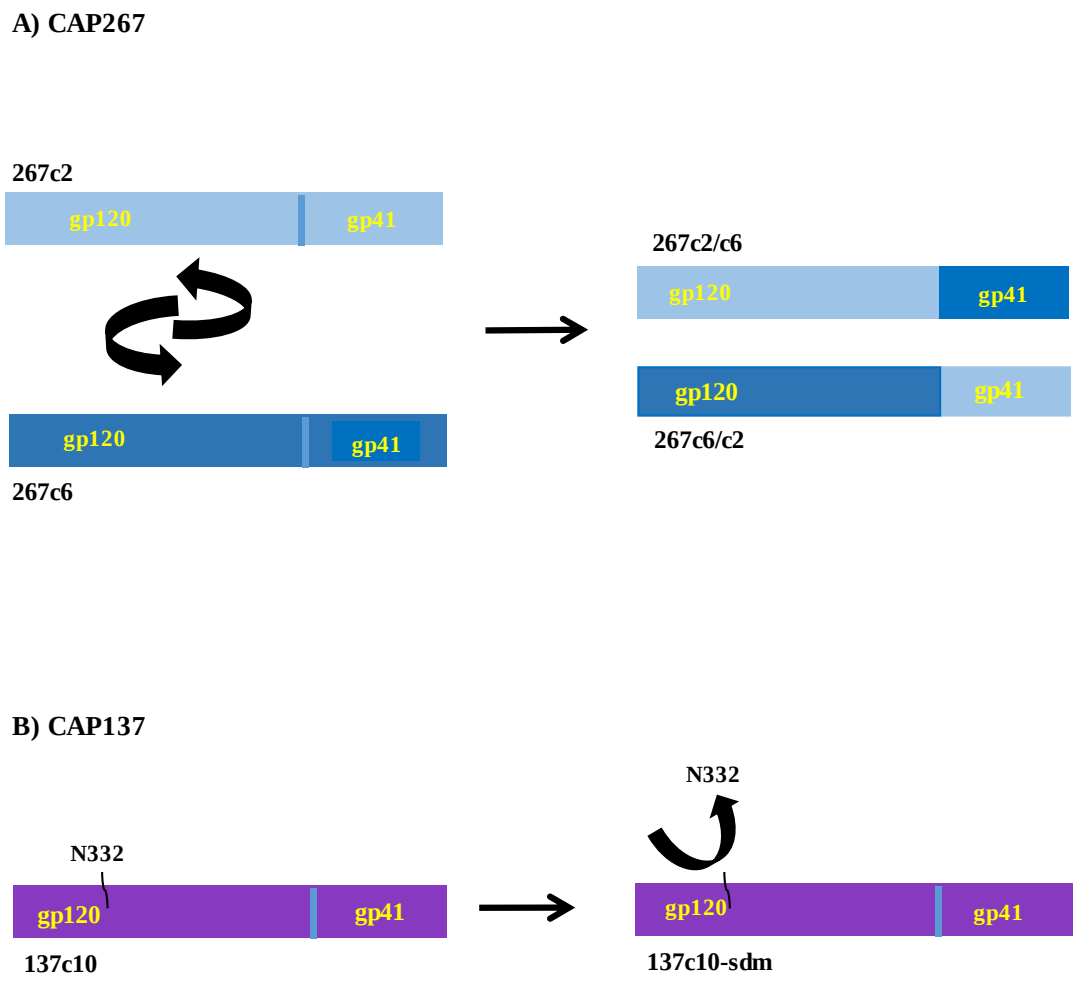


Figure 2.2. Construction of CAP267 chimeras and CAP137 mutant. A) Two chimeras were constructed by swapping gp120 between 267c2 (representing B virus at 0 mpi) and 267c6 (representing B virus at 12 mpi) indicated with light blue and dark blue, respectively. These chimeras were designated as 267c2/c6 and 267c6/c2. B) CAP137 potential N-glycan site at 332 was removed from 137c10 clone, representing AB virus at 12 mpi.

2.4 Statistical analysis

GraphPad Prism 5.0 software (CA, USA) was used to perform all statistical analysis. One-way ANOVA with Bonferroni correction for multiple comparisons was used to compare entry efficiency between all functional clones. Correlation between *in vivo* frequency and *in vitro* entry efficiency was evaluated using the non-parametric Spearman rank correlation test. All p-values were based on a two-sided test. Significance p value < 0.05.

2.5 Results

2.5.1 Diversity of variants infecting dual infected individuals

Longitudinal analysis of full-length SGA-derived *env* sequences of four dual infected CAPRISA 002 participants, CAP37, CAP84, CAP137 and CAP267 characterised viral evolution over the first year of infection. Changes in diversity over time was determined by measuring the mean intra-participant pair-wise DNA distance at each time point as well as diversification from the variant with the highest frequency at 0 mpi (Figure 2.3A and C).

There was an apparent initial peak in viral diversity for all participants; with variants infecting CAP137 having the highest mean DNA distance (6.5 %) at 3 weeks post infection (Figure 2.3). Diversity of variants infecting CAP84, CAP37 and CAP267 seemed to stabilize between 6 and 12 mpi with an average rate of change of 0.035, 0.05 and 0.08 % per month, respectively. Quasispecies' diversity infecting CAP137 however, continued to decrease after the initial peak to 2.2 % at 12 mpi. Mean diversity of variants infecting CAP84 seemed to change much slower than the other participants with diversity increasing from 0.07 % at 0 mpi to 1.7 % at 12 mpi. When the mean DNA distance of total variants at all time points was compared between participants, there was a significant difference between CAP137 and CAP84 ($p < 0.001$) (Figure 2.3B), confirming longitudinal data (Figure 2.3A) that variants infecting CAP137 had evolved much more rapidly within the first 12 months of infection than viruses infecting the other participants although at 12 mpi, all participants were infected with variants of similar diversity (range = 1.7 – 3.2 %).

An alternative measure of HIV evolution is the change in DNA distance relative to a reference variant, usually the founder virus (Abrahams et al., 2009). Except for those infecting CAP137 for the first 6 months, all variants infecting the dual infected individuals diverged from the parent virus at 0 mpi (Figure 2.3C). Overall, divergence and diversification followed a similar pattern as outlined by Shankarappa et al. (1999) for early infection (Shankarappa et al., 1999). We cannot exclude the possibility of sampling bias resulting in detection of only a subset of variants (Salazar-Gonzalez et al., 2008; Troyer et al., 2005). However, we hypothesised that if variants were not detected then they were unlikely to

influence viral loads and might not be essential to determine the role of Env function in *in vivo* viral outgrowth (Arenas et al., 2016).

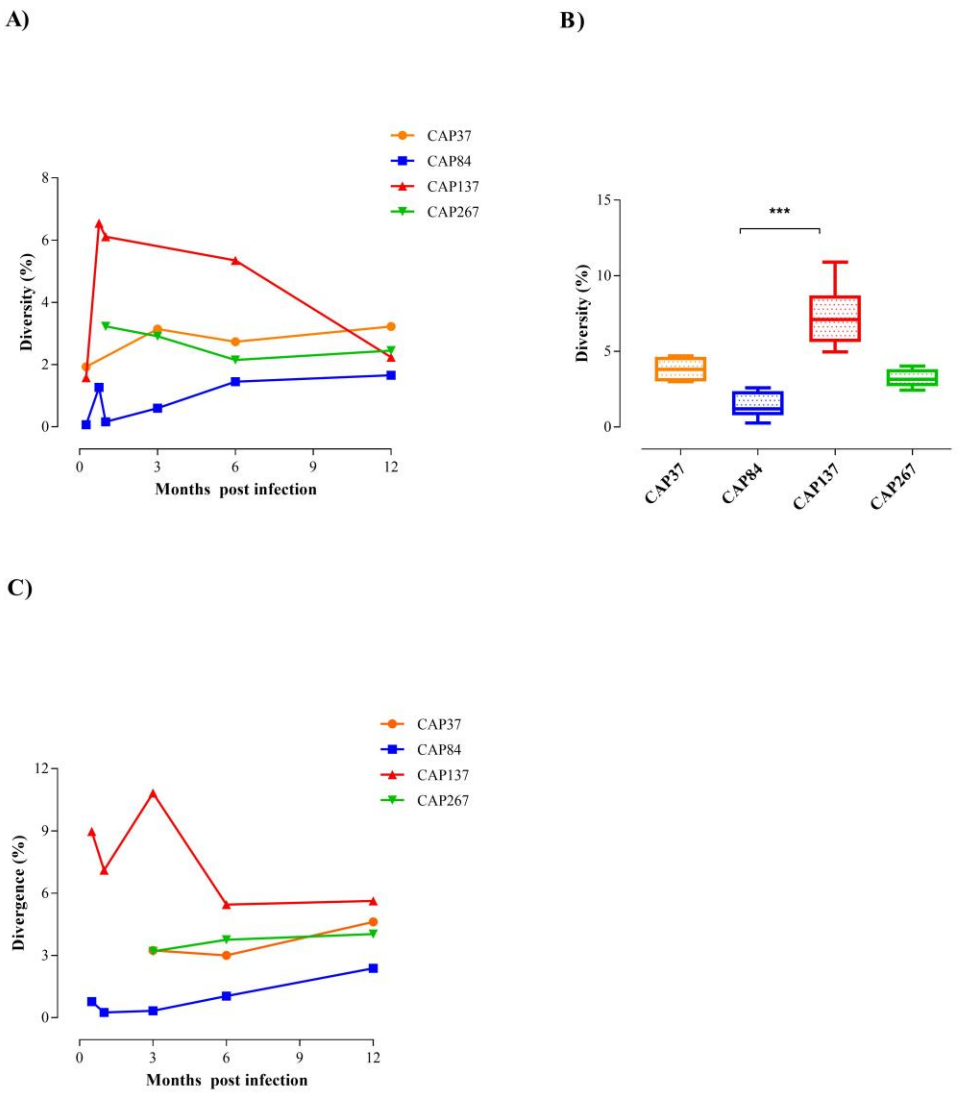


Figure 2.3. Diversity and Diversification of envelope sequence over time. A) Mean pairwise DNA distance (%) between intra-participant *env* sequences generated at 0 mpi, three months, six months and 12 months' post-infection was calculated over time for CAP37 (orange), CAP84 (blue) and CAP137 (red) and CAP267 (green). B) Average pairwise DNA across all time points was compared between participants using Kruskal-Wallis test" with "Dunn's Multiple Comparison Test. (*** < 0.001). C) Divergence was calculated as the difference in DNA distance at each time point relative to the dominant variant present at 0 mpi.

2.5.2 Frequency of viral populations infecting dual-infected individuals

The full-length *env* sequences derived from the four dual infected individuals provided a snapshot of the relative frequency of variants present in the viral population over time: 0, 3, 6 and 12 mpi. We hypothesised that variants at high frequency were most likely driving viral loads in dual infected individuals and would have high *in vivo* replication fitness. Variants were tracked over the course of infection and their relative frequency was indicated as a percentage of total sequences generated at each time point. The relative frequency was used as a surrogate marker of *in vivo* viral fitness (Salazar-Gonzalez et al., 2011).

2.5.2.1 Confirmation of dual infection

Dual infection is defined as infection by variants that group independently from each other, separated by phylogenetically unrelated sequences (Grobler et al., 2004). However, when full-length sequences of CAP37, CAP267 and CAP84 were aligned with phylogenetically unlinked viruses, they did not group separately in Neighbour-joining trees (data not shown). This could be due to transmission of multiple related variants from the donor in a single transmission event, rapid recombination within the recipient and/or the transmission of recombinants derived from the donor. When Neighbour-joining trees of CAP37, CAP84 and CAP267 V3 and gp41 sub-genomic regions were drawn with 30-100 unrelated sequences, two viral populations carried phylogenetically distinct sequences (Appendix B) (Kraft et al., 2012; Pacold et al., 2010; Woodman et al., 2011). This was consistent with the previous finding of Woodman et al. (2011) who first identified these individuals as dual infected by analysing C3-V3 regions using Heteroduplex Mobility assay (HMA) (Woodman et al., 2011). It was highly likely that these variants were recombinants of the initial "parents" and thus shared genetic information.

2.5.3 Identification of phylogenetically distinct master sequences

In order to track variants over the course of infection we first had to designate variants as 'parents' or master A and B *env* sequences. To do this we used three approaches: 1) full-length *env* sampled at 0 mpi and 2-3 mpi for each participant were analysed by Neighbour-joining trees (Figure 2.4). Variants that clustered separately with low bootstrap values were

designated A and B while those that clustered with neither group were classified as AB recombinants; 2) the maximum DNA distance between viruses A and B were determined to identify two sequences that were the least related of all viruses and to verify that the maximum DNA distance of variants designated as AB recombinants were equidistant to virus A and virus B (Table 2.1); and 3) Highlighter plots using A and B parent sequences as masters were used to confirm visually the classification of viruses at early infection (Figure 2.5).

Two viral populations, A and B, grouped separately on the neighbour-joining trees of CAP37 and CAP267 at 0 mpi with maximum DNA distance of 8.3 % and 5.3 %, respectively. Population B infecting CAP37 made up only 5 % of the total population present, whereas CAP267 total viral population at 0 mpi comprised of 29 % B variants (Table 2.1, Figure 2.4). However, CAP84 and CAP137 were infected with homogenous viral populations at the first-time points. A second viral population was detected at 2 mpi for CAP137 and 3 mpi for CAP84. These variants designated B viruses had a maximum DNA distance of 12 % and 3 % from the initial variants at 0 mpi of CAP137 and CAP84, respectively.

Besides populations A and B, two recombinant variants infecting CAP137 were also identified at 2 mpi that not only grouped separately from viruses' A and B but also from each other. Despite differing from one another by 12 %, these two distinct recombinant populations were both designated AB to make it easier to track variants over time. Once variant frequency of AB recombinants was determined, the individual recombined populations were investigated in more detail to determine whether dominance of one recombinant population over another was due to Env entry efficiency. A, B and AB viruses infecting CAP137 had a frequency of 17 %, 50 % and 33 %, respectively at 2 mpi (Table 2.1; Figure 2.4 and Figure 2.5).

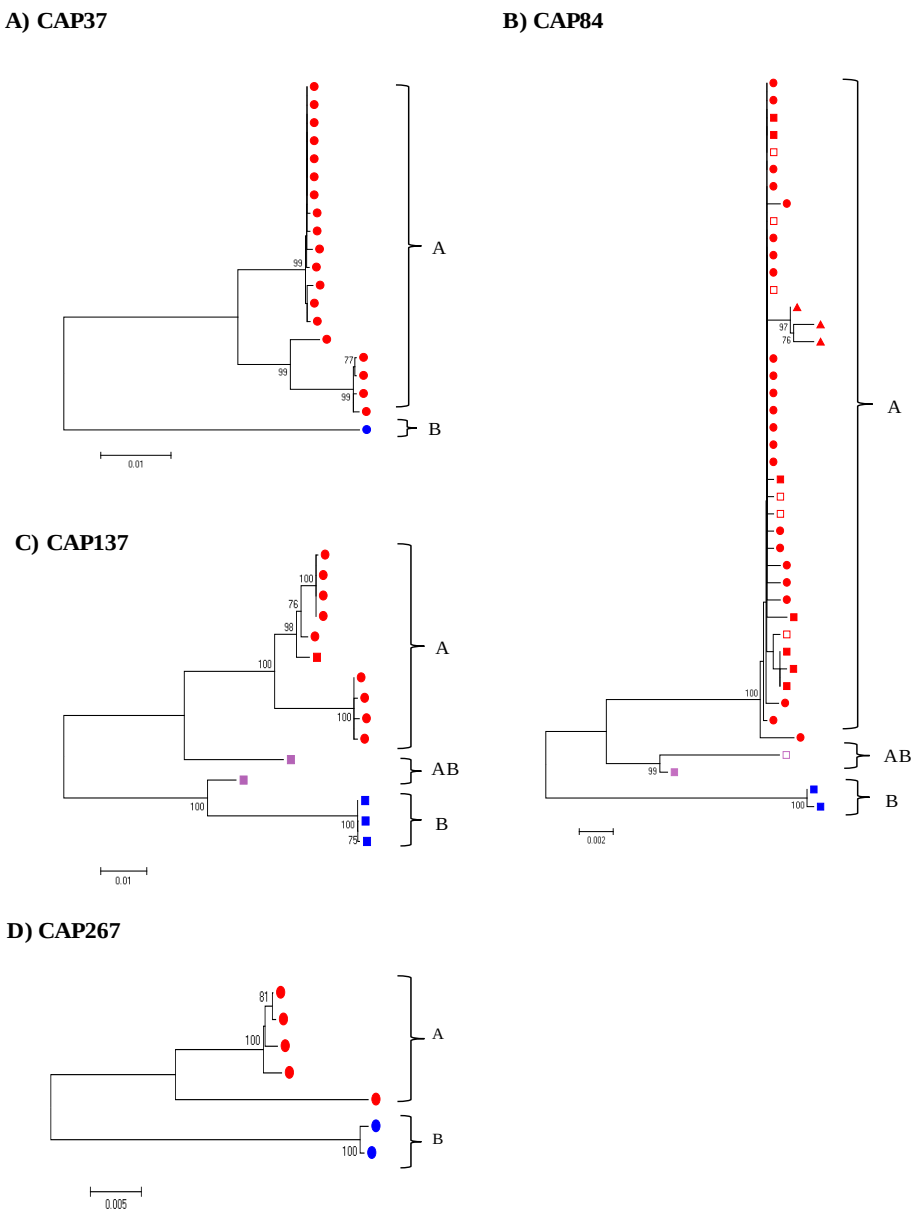


Figure 2.4. Phylogenetic analysis of SGA-derived *env* at early time points. The full-length env sequences of A) CAP37, B) CAP84, C) CAP137 and D) CAP267 from 0 mpi (CAP37, CAP267) or 0 mpi and 2-3 mpi (CAP84 and CAP137) were aligned using BioEdit. Later time points were included for CAP84 and CAP137 as they were infected with a homogenous viral population at 0 mpi and phylogenetically different variants were only detected at later time points. Neighbour-joining trees were used to identify A (red) and B (blue) master sequences with the highest maximum DNA distance and AB variants (in purple) with equidistant maximum DNA distance from both masters. Sequences from 0 mpi are indicated with a circle and variants from later time points are labelled with a square. CAP84 sequences from 0 mpi, 0.5 mpi, 1 mpi, and 3 mpi are indicated with circle, triangle, empty red square and filled square, respectively. Boot strap values lower than 70 % are not shown.

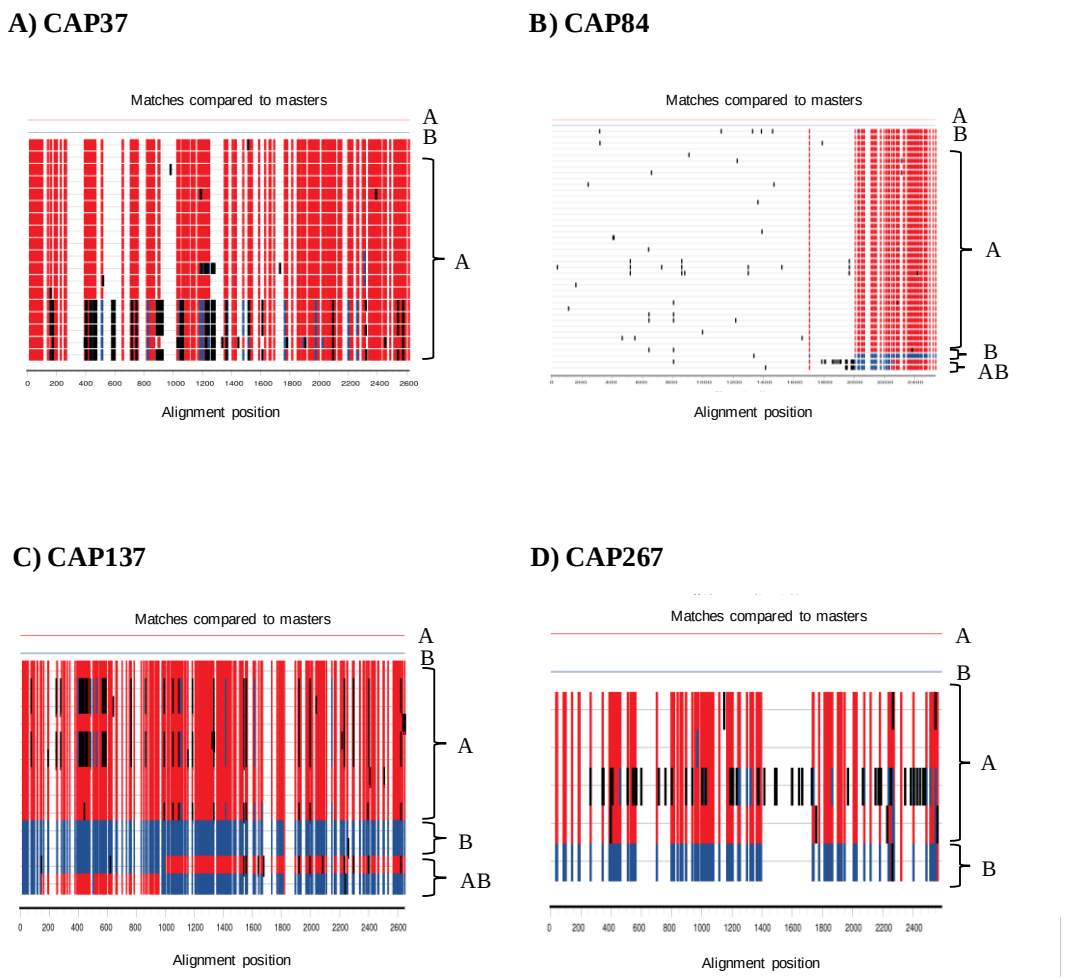


Figure 2.5. Highlighter Plot of sequences generated at early time points. Highlighter Plot of SGA-derived env sequences generated at earlier HIV-positive time points, showing variants from each participant: A) CAP37, B) CAP84, C) CAP137 and D) CAP267. Env sequences with the highest maximum DNA distance which grouped separately on Neighbour-joining trees were used as master sequences (virus A: red and virus B: blue) to which all other sequences were compared. Variants that grouped separately were indicated A and B in CAP37 and CAP267 and variants identified at ~2 mpi for CAP137, and up to 3 mpi for CAP84 were designated A, B and AB. To simplify diversity, all recombinants were labelled AB irrespective of whether they grouped together or not. Red lines represent sequences that match with A and blue lines indicate sequences that match with virus B. Black lines indicate sites that are unique to both A and B variants and empty spaces indicate sequences that are common to both master sequences.

CAP84 was also infected with a homogenous viral population at initial time points until a second variant was detected at 3 mpi that grouped separately from virus A, albeit only within gp41. Despite variation limited to gp41 only, to compare evolution of sequences over time, this variant was designated virus B as it had the highest DNA distance from virus A (3 %). Virus B had a frequency of 20 % at 3 mpi. In addition, two minor sequences (10 % frequency) which appeared to be recombinants of A and B was also detected (Table 2.1, Figure 2.4B and Figure 2.5B).

2.5.3.1 The evolution of variants over the course of infection

Highlighter plots of longitudinal full-length *env* for each participant indicated 'parents' (the initial transmitted viruses that were detected), evolved 'parents' (accumulation of point mutations over time in parent sequences), recombinants (carrying regions of sequence from each parent) at the earliest time point or emergent recombinants (appeared during the course of infection) (Figure 2.6). Identifying which variants grouped with A, B or AB populations was difficult using only Highlighter plots (Figure 2.6) and Neighbour-joining trees (Appendix B). To simplify the complexity and thus determine frequency of variants, pair-wise DNA distance between all sequences and master sequences, A and B were determined. If a variant was more similar to virus A than B, it was classified as virus A and those that were equidistant were designated AB recombinants (Table 2.1). Evidence of recombination was confirmed using RIP analysis (Figure 2.7).

CAP37

Highlighter plots suggested that early recombination occurred between variants infecting CAP37, (Figure 2.6A). Therefore, variants designated virus B were likely recombinants of virus A (or vice versa) and an unidentified phylogenetically distinct parent. Furthermore, full-length *env* sequence analysis indicated that CAP37 viral population A dominated during early infection (Figure 2.6A). However, recombinant variants AB, sharing regions (mainly in signal peptide, V1 and gp41) from both A and B viruses emerged at 6 mpi with a frequency of 17 %, and outgrew all other viruses at 12 mpi with a frequency of 100 % (Table 2.1, Figure 2.6A, and Figure 2.8A). Neighbor-joining trees and Highlighter Plots showed that the AB population at 12 mpi could be divided into two sub-groups that differed by 2.6 % (Figure

2.6). One group was more similar to virus A and thus clustered with A virus, while the other was more closely related to the other parent as it carried C3 and gp41from B virus (Figure 2.7 and Appendix B). These two recombinant populations had similar frequency of 45 % and 55 % at 12 mpi and thus we could not confirm which variant was dominant at this time point.

CAP84

Master sequences A and B differed only in gp41 and viral population B remained minor throughout the first year of infection. As viruses A and B shared gp120 sequence, virus B was likely a recombinant. However, by 12 mpi, virus B had evolved to include regions of virus A gp41 sequence as well as unique changes within the signal peptide (Figure 2.6). This viral population had an average DNA distance of 2.8 % and 2.5% from A and B, respectively, and was designated as an AB variant. AB virus dominated other viral populations with a frequency of 70 % at 12 mpi) (Table 2.1, Figure 2.8), suggesting that evolution in the signal peptide and/or gp41 could have conferred a growth advantage to this population.

CAP137

CAP137 was infected with highly diverse variants as the three viral populations A, B and AB identified by ~2 mpi fluctuated and recombined over the course of infection. Virus A evolved to include C3, V4 and/or gp41 regions from virus B resulting in an emergent AB population at 12 mpi distinct from the initial recombinant identified at 0 mpi. Similar to variants infecting CAP37, the emergent AB population at 12 mpi could be sub-divided into two groups with an average DNA difference of 2.1 % due to differences mainly within gp41 (HR2 and MPER regions). Although, one AB group shared higher similarity with virus A (AB/A) and had a frequency of 31 % and the other AB group was more similar to virus B (AB/B) (Figure 2.7) with a frequency of 69 %, these sequences did not cluster separately and grouped with virus A, albeit with low bootstrap values (Table 2.1, Figure 2.8). The outgrowth of population AB/B suggests that recombination selected for a fitter virus *in vivo* over time through recombination within the C3, V4 and gp41 regions.

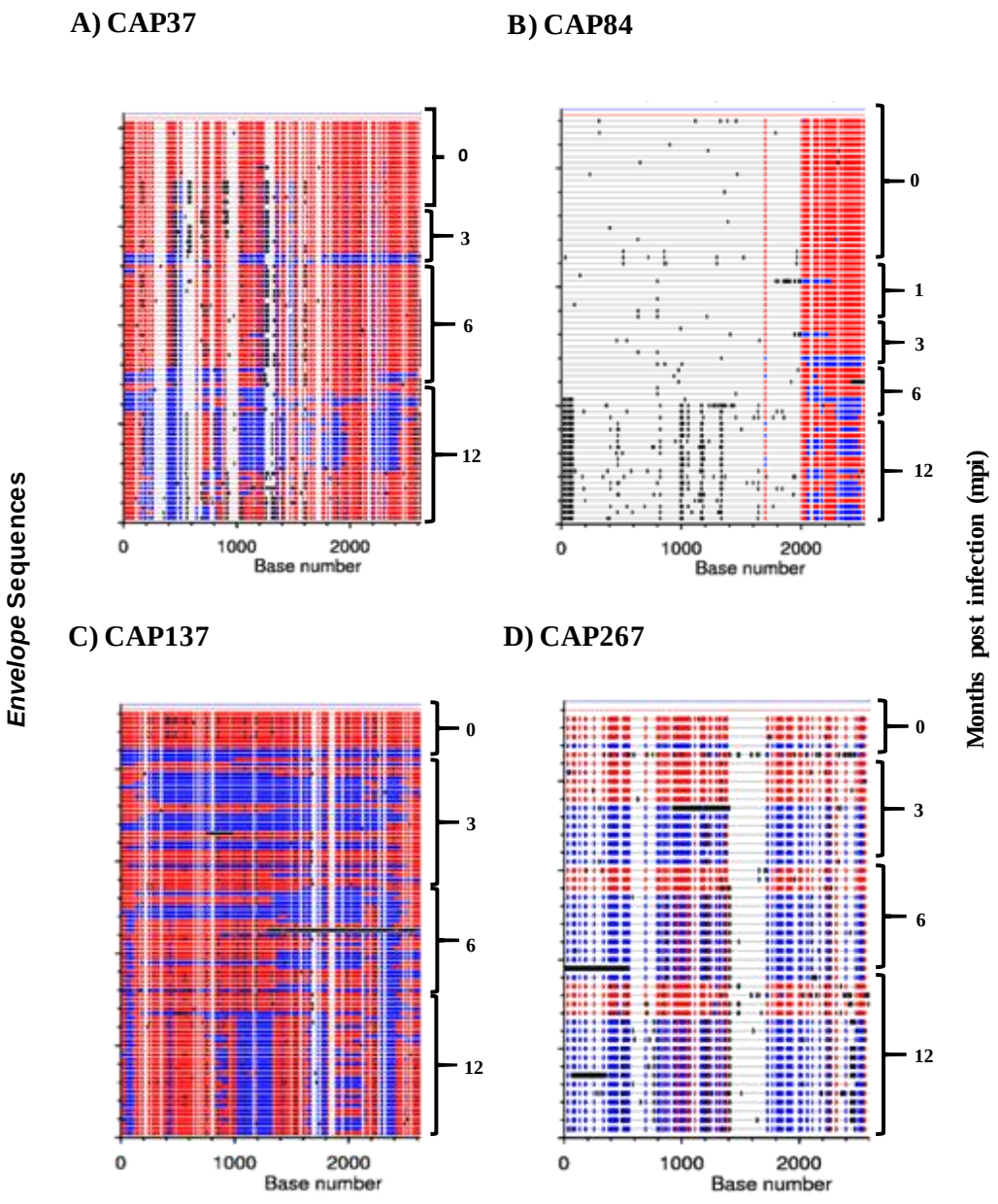


Figure 2.6. Sequence diversity over the course of infection. The full-length env sequence alignments of A) CAP37, B) CAP84, C) CAP137 and D) CAP267 were analysed using Highlighter plot (www.lanl.gov) with sequences from 0 mpi, 3, 6 and 12 mpi compared to A and B masters. A and B viruses are shown by red and blue lines, respectively while black lines indicate unique sequence not present in either master sequence. Sequences identical in both master sequences are not coloured.

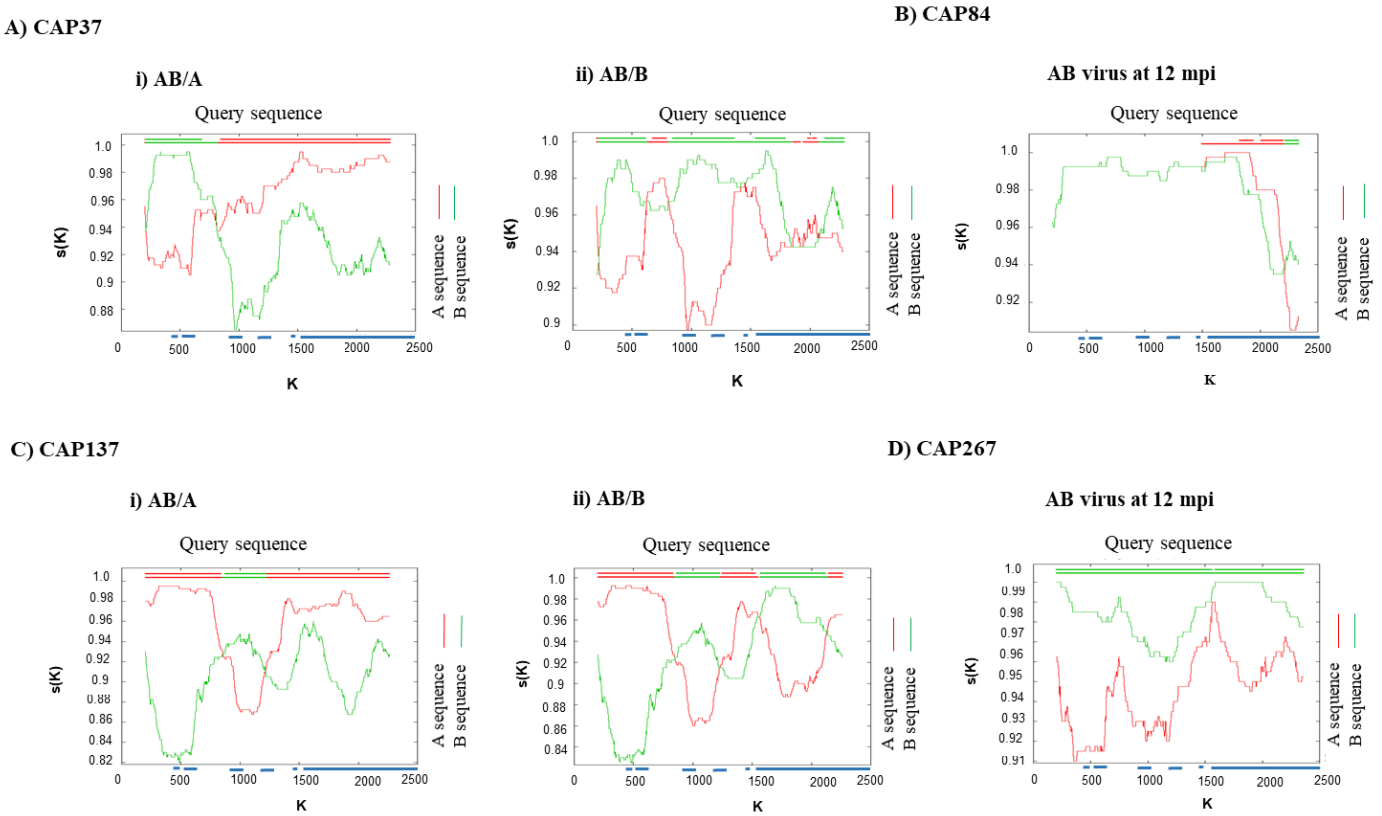


Figure 2.7. RIP analysis to determine whether AB variants are recombinants. A) At 12 mpi CAP37 AB/A sub-population was more similar to virus i) A and AB/B variant was more similar to virus ii) B due to additional C3 and gp41 virus B regions. B) CAP84 AB env at 12 mpi carried regions in gp41 that were of both A and B viruses. C) At 12 mpi CAP137 AB viral population could be separated into two sub-groups with i) AB/A more similar to virus A carrying C3 and V4 from virus B, and ii) AB/B was more similar to virus B as it carried both C3, V4 and gp41 of this variant. D) CAP267 viral population B at 12 mpi was similar to virus B only. Virus A is shown in red and virus B is indicated in green and the top line is the query sequence. The x-axis (k) represents the query sequence position at the centre of the moving window of 400 bp. The y-axis, s(k), shows the similarity between that window of query sequence and virus A and virus B. Relative positions of variable loops indicated in blue lines along x-axis.

Table 2.1. Characterisation of viral populations over time in dual infected individuals

Participant ID	[a]Visit	Time post-infection (weeks) [b](months)	Average Population Diversity (%) from master	Max DNA distance from A/ Max DNA distance from B (%)	[c]Number of groups	[d]Group ID	Frequency (%)	[e]Clone ID
CAP37	1_18	2 (0)	0	0.00/8.3	2	A	95	37c1
				2.91/8.97		A		37c2
				8.25/0.00		B	5	37c3
	2_05	13 (3)	3.24		2	A	82	N/A
						B	18	N/A
	3_08	21 (6)	3.00	2.24/8.4	3	A	73	37c4
				5.5/3.1		B	10	37c5
				4.9/4.4		AB	17	N/A
	3_16	56 (12)	4.62	3.7/7.10	2	AB/A	45	37c6
				5.4/3.9		AB/B	55	37c7
				5.4/3.9		AB/B	55	37c8
CAP84	1_00	1 (0)	0	0.00/2.8	1	A	100	84c1
	[f]2_00	3 (0.75)	0.78		1	A	100	N/A
	2_01	4 (1)	0.26	0.00/2.8	2	A	86	N/A
				1.9/2.9		AB	14	N/A
	2_05	10 (3)	0.34	0.00/2.8	3	A	70	N/A
				2.84/0.00		B	20	84c2
				1.3/2.3		AB	10	NA

	3_08	19 (6)	1.04		2	A	71	N/A
						B	29	N/A
	3_16	54 (12)	2.38	1.07/3.84	2	A	30	84c3
				2.53/2.37		AB	70	84c4
CAP137	1_13	2 (0)		0.00/0.00	1	A	100	137c1
				2.72/12.85		A	100	137c2
	[f]2_00	7 (2)	8.97	0.7/12.2	3	A	17	N/A
				12.24/0.00		B	50	137c3
				5.2/7.9		AB	33	N/A
	2_04	12 (3)	7.10	0.86/12.68	3	A	34	137c4
				10.59/1.59		B	48	137c5
						AB	17	N/A
	2_05	14	10.82		1	B	100	N/A
	3_08	23 (6)	5.45	4.50/8.45	3	[g]A	65	137c6
				10.31/2.28		B	22	137c7
				5.10/7.50		AB	13	137c8
	3_15	52 (12)	5.63	3.8/9.2	2	AB/A	31	137c9
				5.62/7.56		AB/B	69	137c10
CAP267	[f]2_01	6 (0)		0.00/5.32	3	A	71	267c1
				5.32/0.00		B	29	267c2
	2_04	10 (3)	3.20	0.19/5.36	2	A	42	267c3
				5.35/0.63		B	58	267c4

	3_08	20 (6)	3.76		2	A	23	N/A
						B	77	N/A
	3_15	52 (12)	4.03	1.13/5.48	2	A	24	267c5
				5.39/1.4		B	76	267c6

a Time of sampling is phase of study_visit number

b To simplify the comparison between sequences from different participants, the visit from which we got the first sample was classified as 0 months post-infection (mpi), 3, 6 and 12 mpi even though number of weeks at sampling differs slightly between participants.

c The number of viral populations that grouped separately according to Neighbour-joining trees or had different similarities to A and B viruses based on max DNA distance

d Initial, phylogenetically distinct viral populations were designated A and B and variants that emerged after 0 mpi were classified as AB if they differed equally from both A and B viruses. Some AB populations comprised sub-groups where one was more similar to virus A and one was more similar to virus B and this is indicated with /A or /B, respectively.

e Clone ID: Participant ID number clone name in chronological order for each participant

f Samples from enrolment to the CAPRISA002 study.

g Although 137c6 clustered with virus AB, the DNA distance analysis indicated that it was more closely related to virus A.

CAP267

Viral populations A and B were identified at 0 mpi with a max DNA distance of 5.3 %. Virus A dominated the quasispecies with a frequency of 71 %. Although, viral population B was a minor variant at 0 mpi, it evolved over time with changes mostly in gp41 and became the major population at 12 mpi with a frequency of 76 % (Figure 2.6, Figure 2.8 and Table 2.1). Although there was no clear evidence of recombination between A and B viruses over the course of infection (Figure 2.7), it is highly likely that virus B was a recombinant of A (or A was a recombinant of B), suggesting that recombination occurred early within this participant (Figure 2.6) and that there was apparently no further recombination between the two viruses over 12 months of infection (Figure 2.7), most likely because further events did not confer a growth advantage to the virus (Arenas et al., 2016).

At a glance, recombination, common to most participants, occurred mainly within C3 (n=2) and gp41 (n=3) although a number of other regions also recombined (Figure 2.7). Overall, even though viruses infecting dual infected individuals evolved along dissimilar pathways, recombinant variants tended to dominate the population by 12 mpi (Figure 2.6). CAP137 AB variants that emerged (or were detected) at 12 mpi had the highest frequency one year after infection. Variant B infecting CAP37, CAP84 and CAP267 were likely recombinants of virus A and an unidentified parent and thus recombination had probably occurred close to 0 mpi or within the donor. Virus B of CAP267 and AB, viruses generated through further recombination between viruses A and B, infecting CAP37, CAP84 and CAP137 became the dominant viruses at 12 mpi. This suggests that recombination plays an important role in the generation of variants able to out-compete other viruses.

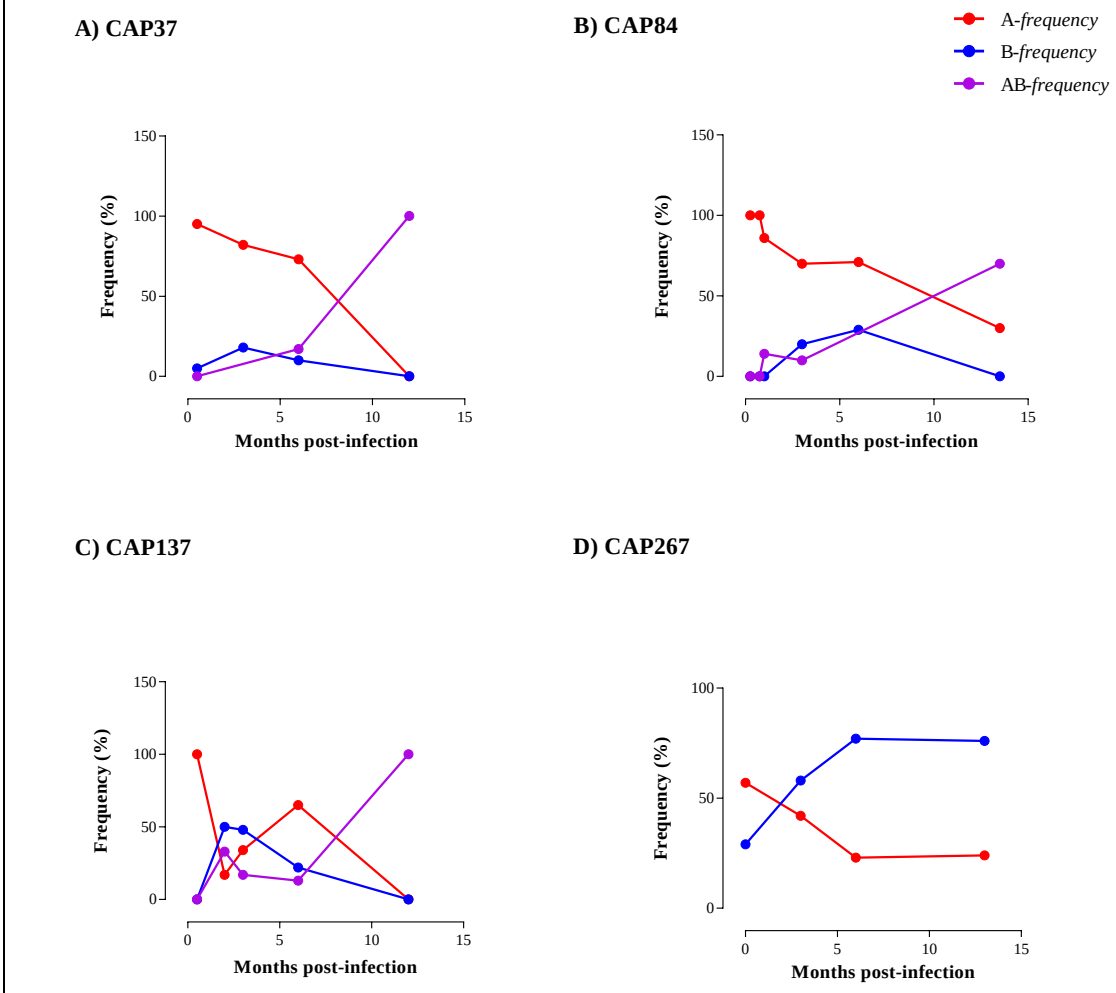


Figure 2.8. In vivo frequency of viral populations over time. The *in vivo* frequency of viral populations over time in each participant: A) CAP37, B) CAP84, C) CAP137 and D) CAP267 showing change in frequency of viruses A (red), B (blue) and AB (purples) and outgrowth at 12 mpi. *In vivo* frequency (%) of variants was determined by calculating the relative proportion of sequences at each time point for each participant. Viruses not detected at 0 mpi had 0 % frequency.

2.5.4 Changes in Env Entry efficiency of viruses infecting dual infected individuals over time

We showed previously that Env entry varied over the course of HIV-1 dual infection (Gordon et al., 2016). Furthermore, it has been shown that Env diversity via mutation and recombination influenced *in vitro* HIV-1 fitness (Arenas et al., 2016; Gordon et al., 2016; Haddox et al., 2016). In order to investigate whether changes in Env entry efficiency was influenced by diversification, SGA-derived *env* amplicons, representing viral populations, A, B and AB were cloned into a mammalian expression system. PSVs were generated and entry efficiency was compared between clones using the TZM-bl cell line. To compare entry efficiency of PSVs over the course of infection, across independent experiments, all measurements were calculated relative to virus A at 0 mpi.

2.5.4.1 Association between pseudovirion entry efficiency and Env evolution

CAP37

Eight functional Env clones representing viruses A, B and AB were generated: three clones from 0 mpi, two clones from 6 mpi and three clones from 12 mpi. The three 12 mpi clones represented the two sub-groups of AB virus, with two from population AB/B (37c7 and 37c8) (Table 2.1; Appendix B). CAP37 was infected with highly diverse variants and this was reflected in the variation in entry efficiency between PSVs (range = 23 – 141 %). PSVs representing virus A had 0.5-fold lower entry efficiency than B-associated PSVs at 0 mpi although this difference was not significant. At 6 mpi virus A entry efficiency did not change but there was an apparent decrease in virus B PSV entry. At 12 mpi there was no significant increase in entry efficiency of all AB recombinants compared to PSVs of virus A and B at previous time points. Despite extensive recombination within AB/B, PSVs had the lowest entry efficiency suggesting that recombination was associated with a fitness cost. On the contrary, the AB/A subgroup had significantly higher entry efficiency compared to AB/B PSV ($p < 0.001$), suggesting that some recombination events over the course of infection had selected for variants with higher Env entry efficiency.

CAP84

Due to the low viral diversity of variants infecting CAP84 only four functional Env clones were generated representing viral population A at 0 mpi (84c1), virus B at 3 mpi (84c2) and viruses A and AB at 12 mpi (84c3; 84c4). PSV entry efficiency of virus A was significantly higher than that of B-associated PSV ($p = 0.01$), although at 12 mpi, virus A PSV entry had significantly declined, suggesting that diversification within population A had come with a fitness cost (Figure 2.9). Comparison of Env sequences between clones representing virus A at 0 mpi and 12 mpi indicated that they differed in 39 amino acids, with most changes occurring within the gp41, suggesting that these changes were responsible for the decrease in entry efficiency of virus A over time. At 12 mpi virus A and virus AB Env sequences were nearly identical except for point mutations within gp120 and recombination within gp41 that incorporated regions of virus B (Figure 2.6). As virus AB had the same entry efficiency as PSV A sampled at 0 mpi, and significantly higher entry efficiency than PSV A at 12 mpi, introduction of virus B gp41 sequence apparently rescued Env fitness. Therefore, diversification of virus A over the course of infection led to a less fit variant but recombination within gp41 might have restored fitness.

CAP137

A total of ten Env clones were generated representing variants infecting CAP137 over time. Two clones represented viral population A at 0 mpi, one representing virus B at ~2 mpi, two representing A and B viruses at 3 mpi, three representing A, B and AB at 6 mpi, and two representing viral population AB at 12 mpi (Table 2.1). The variation in entry efficiency between the clones was high, ranging from 10 % to 239 %. There was no significant difference in the entry efficiency between A and B PSVs at all time points and their entry efficiency decreased over 12 months of infection, suggesting that diversification had reduced Env fitness. However, further recombination between A and B viruses in C3, V4 and gp41 regions resulted in the emergence of recombinants at 12 mpi with significantly higher PSV entry than virus A at 0 mpi and virus B at 2 mpi ($p < 0.001$) (Figure 2.9). Viral population AB at 12 mpi comprised of two subgroups with one, AB/B, carrying additional gp41 sequences from virus B and having significantly higher entry efficiency than all other Env

clones except AB/A. This could suggest that sequential recombination events, incorporating virus B regions, first in gp120 (C3 and V4 regions) and then in gp41 rescued the loss of virus A Env fitness.

CAP267

To compare the entry efficiency of viruses from CAP267, six of the SGA-derived *env* sequences were cloned with two clones representing A and B viral populations at 0 mpi, 3 mpi and 12 mpi. At 0 mpi, PSV generated from viral population A had higher entry efficiency than PSV B although this difference did not reach significance. However, diversification of B virus over time was associated with a significant increase in entry efficiency as PSV B at 12 mpi was 4.7-fold more efficient at entering TZM-bl cells than that of PSV B at 0 mpi. Unlike, the other participants, there was no apparent recombination between viruses A and B over 12 months of infection (Figure 2.7) and it seemed as though B variants evolved to higher entry efficiency through point mutations.

2.5.4.2 Selection of Env with high pseudovirion entry efficiency

This study hypothesized that high diversity in dual-infections would facilitate the emergence of fitter recombinant viruses. Highlighter plots and phylogenetic analysis suggested that CAP37, CAP84 and CAP267 were infected with recombinants at or soon after enrolment in the study. However, without the identification of phylogenetically distinct parents, we were unable to confirm that early recombination had occurred. We were able to confirm that additional recombination events occurred between virus A and virus B for CAP37, CAP84 and CAP137 and of these, CAP137 was infected with recombinants at 12 mpi with increased entry efficiency compared to virus A and virus B at early time points. Furthermore, recombinant variants infecting CAP37 and CAP84 at 12 mpi had significantly higher entry efficiency than other viruses detected at the same time point. Therefore, recombination within the first year of infection seemed to contribute to improved Env entry efficiency of variants 12 mpi. However, there was no overall correlation between maximum Env diversity at each time point and corresponding entry efficiency of the dominant Env at that time point (data not shown) ($p = 0.08$). This suggests that high diversity of populations does not translate

into the emergence of variants with increased Env entry efficiency, but that rapid selection of recombinants conferred a survival advantage.

Despite fluctuations in entry efficiency within time points and over the course of infection, at 12 mpi, CAP37, CAP84, CAP137 and CAP267 were consistently infected with variants that had higher entry efficiency than other variants at the same time point. Escape from antibody and CTL immune responses occurs mostly during early infection characterised initially by rapid shuffling and toggling of polymorphisms within hypervariable regions and CTL epitopes until fixation of escape mutations and slowing of diversification (Abrahams et al., 2013). This is apparent in the initial high diversity of variants followed by convergence and plateauing of diversity and divergence for all four dual infected individuals at 12 mpi (Figure 2.3). Some CTL epitopes previously linked to disease progression and targeted by HLA alleles of CAP37, CAP84, CAP137 and CAP267 were analysed for emergence of polymorphisms over the course of infection (Appendix B). The analysis indicated fixation of putative escape mutations by 12 mpi. Thus, during the first months of HIV infection, there could be rapid sampling of polymorphisms forming a heterogeneous population with varied abilities to escape from immune responses with random Env entry efficiency. Ultimately, however, fixation of mutations that not only allows immune escape but also confers Env phenotypic traits that best facilitate viral survival are selected, resulting in the selection of variants with high viral fitness. Therefore, it is interesting to speculate that the entry efficiency of variants fluctuated seemingly randomly during the first months of infection until immune escape and optimal replicative fitness was reached.

2.5.4.3 Association between Entry efficiency and variant frequency

This study hypothesised that viral outgrowth in dual infected individuals was due to enhanced Env entry efficiency. In order to evaluate the relationship between the *in vivo* outgrowth of variants in dual infected individuals and *in vitro* Env entry efficiency, we compared the entry efficiency of clones representing populations A, B and AB to the relative frequency of each viral population per time point.

CAP37

Virus A was the dominant population at 0 mpi, but PSV entry efficiency was lower than virus B although this difference was not significant (Figure 2.9A). Furthermore, despite emergent AB/B recombinants having marginally higher frequency than virus AB/A at 12 mpi, it had significantly lower entry efficiency ($p < 0.001$), suggesting that viral outgrowth in CAP37 was not due to Env entry efficiency. However, at the population level, the outgrowth of recombinants tended to coincide with enhanced average Env entry efficiency of AB viruses (Figure 2.10A). The discrepancy between sub-population frequency and PSV entry efficiency could be due to the Env clones not representing the viral fitness of the population. Alternatively, fluctuating frequency of viral populations during early infection could be due to immune escape masking a clear relationship between viral outgrowth and Env entry efficiency. Interestingly, CTL escape mutations were identified approximately 12 wpi (Appendix B) whereas Abrahams et al, (2013) showed that CTL escape in Env occurred 2-5 wpi. The authors suggested that immune responses were the driver of early HIV-1 diversification in mono-infections *in vivo* and the introduction of escape mutations could randomly influence Env entry efficiency (Abrahams et al., 2013). This could suggest that CTL escape occurs earlier in mono-infections than dual infections. However, further analysis of the dual-infection sequences needs to be carried out to confirm this suggestion and whether recombination plays a role in selection of CTL escape.

CAP84

The entry efficiency of PSV B was significantly lower than PSV A ($p < 0.01$) at 0 mpi which correlated with variant frequency. At 12 mpi, virus A entry efficiency was significantly reduced compared to enrolment and this coincided with a drop in frequency to 30 % (Figure 2.9B). Whereas virus A seemed to lose fitness, virus AB had significantly higher entry efficiency and frequency than virus A at 12 mpi. Therefore, there was an association between entry efficiency and frequency at each time point with the recombinant having the highest frequency and entry efficiency relative to virus A (Figure 2.9B). This trend was confirmed at the population level when the entry efficiency was compared to variant frequency (Figure 2.10B). As Env representing virus A and the AB recombinant at 12 mpi differed due to

changes within gp41, it is possible that virus B gp41 contributed to enhanced Env entry efficiency of AB.

CAP137

While PSV B entry efficiency varied according to its population *in vivo* frequency over time, the decrease in entry efficiency of PSV A did not correlate with its *in vivo* frequency. At 6 mpi virus A made up 65 % of the viral population but the entry efficiency of PSV A was similar to that of the other variants at that time point. Similar to CAP37, PSV entry efficiency did not always track changes in population frequency during early infection possibly due to the influence of CTL escape mutations on Env fitness (Abrahams et al., 2013) noting that sequences as early as 2-3 mpi showed putative escape within the HLA B*5802 epitope HF9 (Appendix B). However, entry efficiency could be impacting outgrowth of the recombinants as the average entry efficiency of AB/A and AB/B coincided with change in population frequency (Figure 2.10C). Further, analysis of individual clones at 12 mpi indicated that PSVs representing the AB/B sub-group had marginally higher entry efficiency than AB/A that could be responsible for the 2–fold higher *in vivo* frequency (Figure 2.9C). The genotypic difference between AB/A and AB/B is the presence of additional virus B gp41 sequence, suggesting that changes in gp41 might have contributed to enhanced entry efficiency and *in vivo* outgrowth.

CAP267

Viruses A and B entry efficiency fluctuated over time and not all changes corresponded with changes in *in vivo* frequency. There was no evidence of putative CTL escape to explain the lack of association as there was no evidence of polymorphisms within and adjacent to the HLA B*5802 epitope, HF9 (Appendix B). However, at 12 mpi virus B dominated with a frequency of 76 % and had significantly higher entry efficiency than all other Envs (Figure 2.9D). This trend was also apparent when PSV entry efficiency was plotted alongside population frequency (Figure 2.10D). This could suggest that virus B evolved into a variant with enhanced entry efficiency which then contributed to *in vivo* outgrowth.

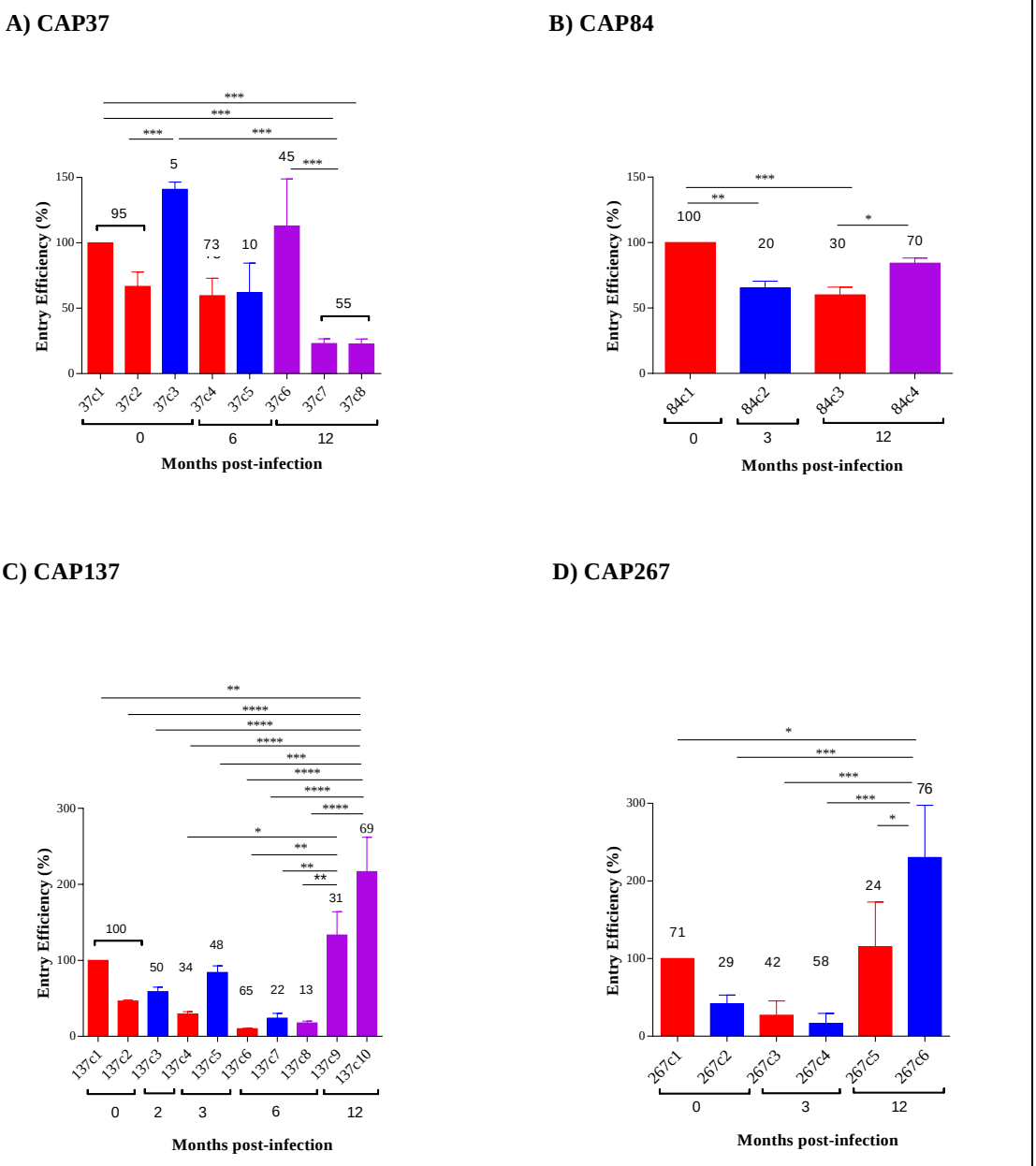


Figure 2.9. Pseudovirion Entry Efficiency over 12 months of infection. The entry efficiency of pseudovirus (PSV) representing virus A (red), B (blue) and AB (purple) infecting A) CAP37, B) CAP84, C) CAP137, and D) CAP267 was compared over time (0, 3, 6 and 12 mpi). TZM-bl cells were infected with PSV equivalent to 100 ng/ml p24 and entry efficiency was measured as a percentage (%) relative luminescent units (RLU) to PSV A at 0 mpi after subtracting background: PSV lacking Env. The in vivo frequency (%) of each virus is indicated at the top of each bar. Results represent three independent biological repeats with error bars indicating standard deviation. One-way ANOVA with Bonferroni correction for multiple comparisons was used for statistical analysis ($p < 0.05$: *, $p < 0.01$: **, $p < 0.001$: ***, and $p < 0.000$: ****).

Of the four participants, three were infected with variants at 12 mpi with high *in vivo* frequency and increased entry efficiency, suggesting that Env fitness was enabling some variants to outcompete others. Overall, there seemed to be a tendency for recombinant variants with high average entry efficiency to have high relative frequency at 12 mpi (Figure 2.10) although correlation analysis showed that this relationship was not significant across all time points ($p = 0.37$, $r = 0.19$) (data not shown). Similarly, populations that dominated at each time point, whether A, B or AB, did not always have the highest entry efficiency ($p = 0.2$, $r = 0.38$) (data not shown). This apparent lack of association between frequency and Env entry efficiency could be due to the dependence of variant survival and outgrowth on immune escape during early infection rather than Env entry efficiency (Abrahams et al., 2013). Finally, for 2/4 participants Env with higher entry efficiency seemed to have recombined within gp41, suggesting that gp41 could play an important role in determining Env fitness and thus virus outgrowth *in vivo*.

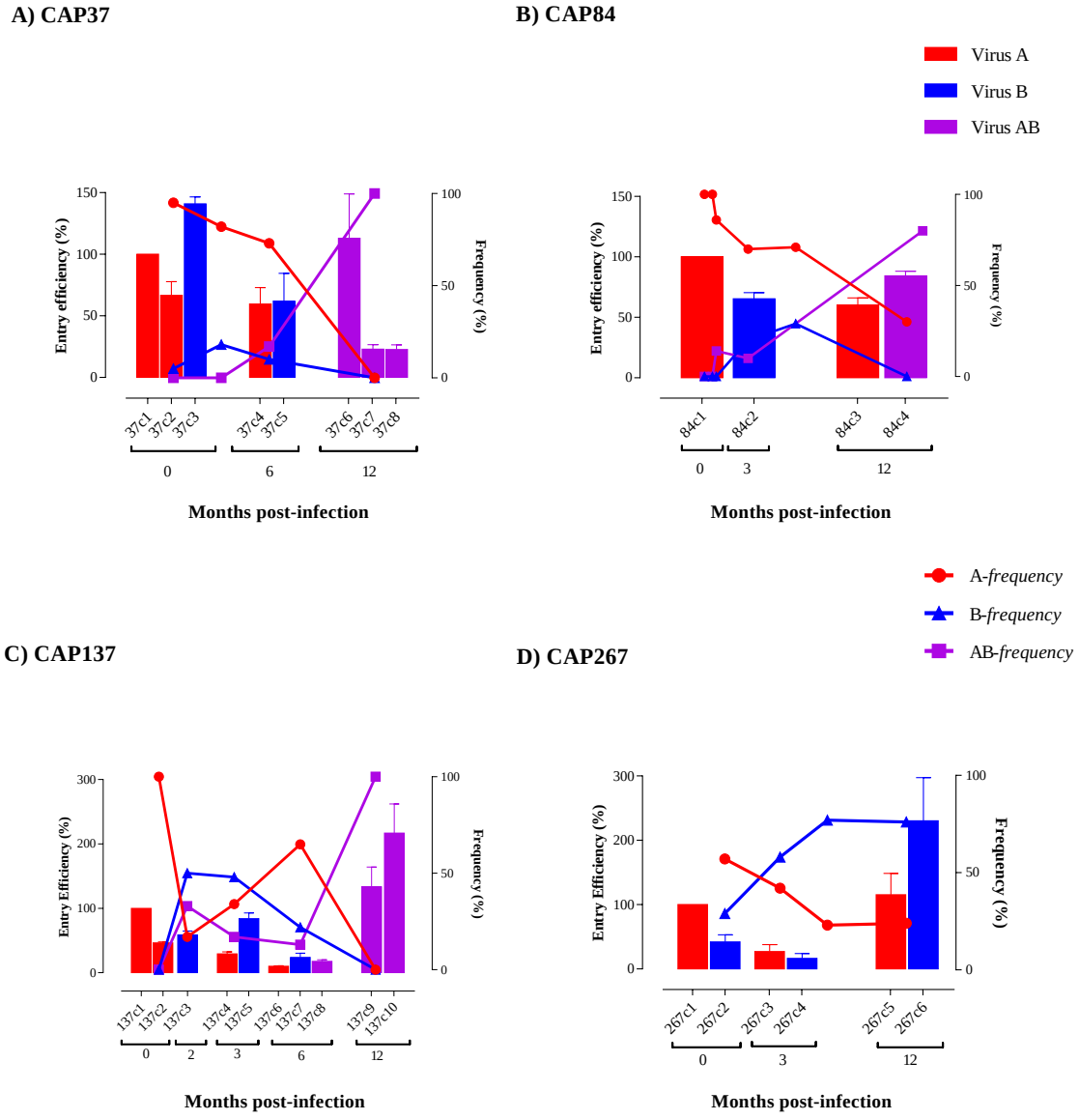


Figure 2.10. Overall relationship between "*in vivo*" frequency and Env entry efficiency over time. Changes in in vivo population frequency (right y-axis) and PSV entry efficiency (left y-axis) of variants infecting A) CAP37, B) CAP84, C) CAP137 and D) CAP267 are shown over months post infection (mpi). *In vivo* frequency (%) of populations was determined by calculating the proportion of sequences relative to total sequences at each time point. The frequency of AB subgroups was not included. PSV entry efficiency is indicated relative (%) to PSV A at 0 mpi. The frequency of A, B, and AB viruses are indicated in red, blue and purple lines, respectively. Entry efficiency of A, B, and AB viruses are indicated in red, blue and purple bars, respectively.

2.5.4.4 Association between Entry efficiency and disease progression

Studies have shown that increased Env entry efficiency is associated with both an increase in viral load (VL) and/or a decrease in CD4+ T cell count (Gordon et al., 2016; Lassen et al., 2009; Mohri et al., 2015; Troyer et al., 2005). Therefore, we determined whether the Env entry efficiency of viral populations A, B and AB infecting dual infected individuals impacted disease progression and whether the entry efficiency of the dominant virus was a determining factor.

For all participants, peak viremia was apparent at the earliest time point and for CAP37, CAP84 and CAP137, the increased VL coincided with the highest entry efficiency of 'parent' viruses A and B compared to later time points. There was a similar trend between VL and PSV entry efficiency across all time points in general, with high VL at 12 mpi coincident with high entry efficiency of AB recombinants (Figure 2.11A-C). Entry efficiency of viruses infecting CAP267 was not linked with VL during the first-year of infection: when Env entry efficiency was low, VL was high and vice versa (Figure 2.11D). However, the increased entry efficiency of CAP267 viruses at 12 mpi coincided with a decrease in CD4 count (Figure 2.11D). This relationship was similar for the other participants where CD4 count levels either increased or plateaued at initial time points before dropping at 12 mpi, concomitant with an increase in entry efficiency of AB viral populations (Figure 2.11A-C).

Despite the apparent association between changes in PSV entry efficiency and VL in 3/4 participants (Figure 2.11) there was no significant correlation between VL and average entry efficiency of all variants at each time point. When a similar analysis was carried out comparing VL to the entry efficiency of the dominant variant only, the relationship still did not reach significance. However, there was a correlation of CD4+ T cell count loss with increased average PSV entry efficiency ($p = 0.046$, $r = -0.59$) (Figure 2.12A) and this relationship strengthened when CD4 count was compared to the entry efficiency of the dominant variant only ($p = 0.02$, $r = -0.71$) (Figure 2.12B). Therefore, there seems to be an overall association between Env entry efficiency and disease progression with three of the

study participants progressing to CD4 levels associated with AIDS-defining illness within the first year of infection (CAP37, CAP137 and CAP267).

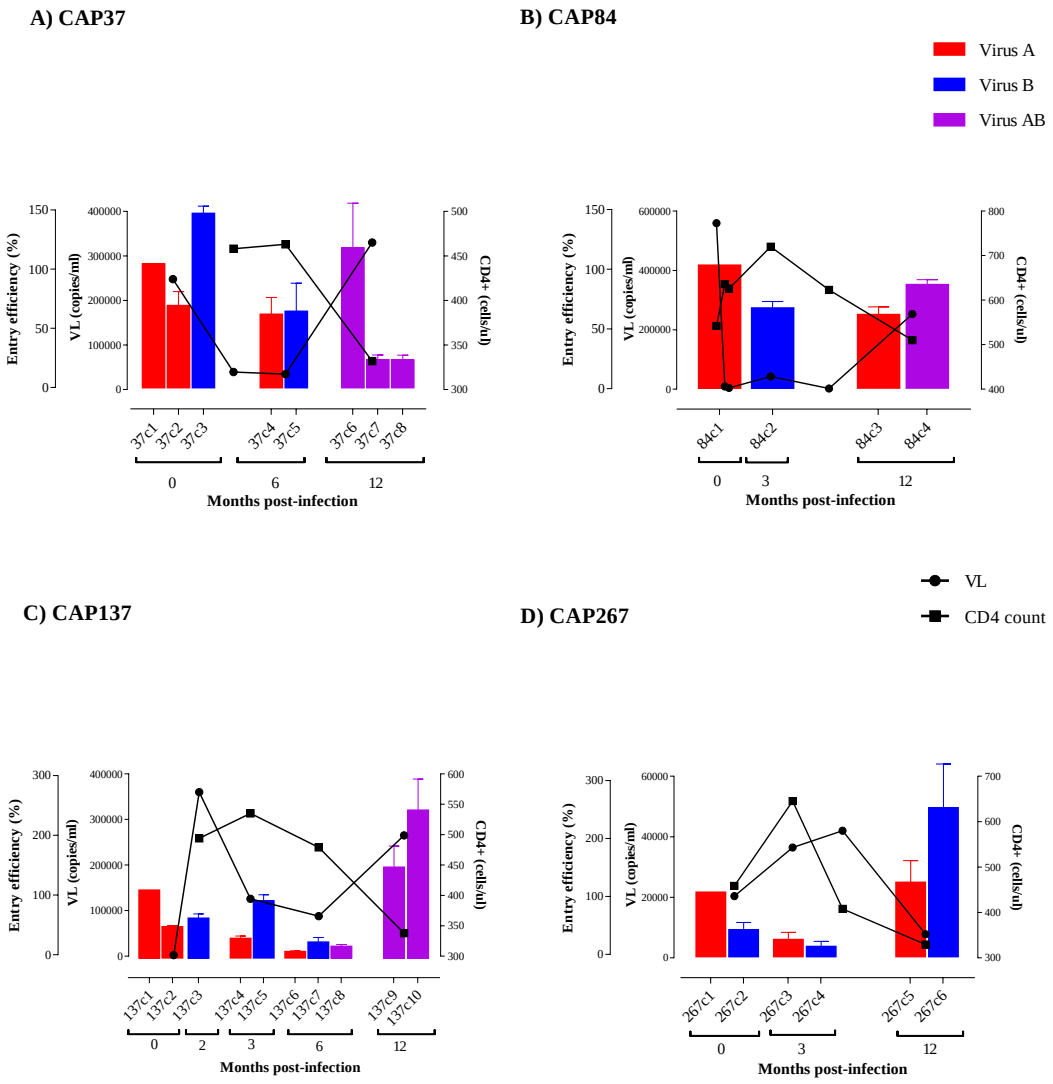


Figure 2.11. Env entry efficiency and disease markers (viral load and CD4 count) over time. Entry efficiency and markers of disease progression, viral load (VL) and CD4+ T cell count were plotted over time. Entry efficiency of PSV A, B and AB is indicated in red, blue and purple, respectively. PSV entry efficiency of all clones was normalised to the entry efficiency of virus A at 0 mpi and indicated as percentage. Changes in VL and CD4 count decline over months post infection (mpi) are shown by black circles and squares, respectively. VL and CD4 counts were included from all timepoints even those when no Env had been cloned or analysed.

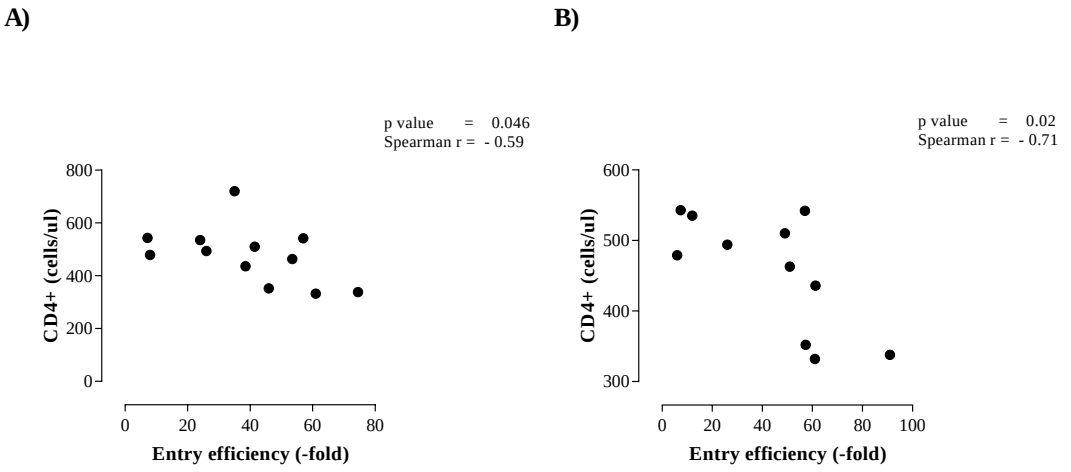


Figure 2.12. Association between pseudovirion entry efficiency and CD4+ T cell count. The entry efficiency of pseudovirus from independent experiments across participants was normalised to the background control and indicated as fold higher than background (-fold). A) The entry efficiency of viruses at each time point per participant, irrespective of population, was averaged and compared with CD4+ T cell count. B) To determine whether the entry efficiency of the viral population dominant at each time point was influencing CD4 counts, the entry efficiency of the virus with the highest frequency was compared to CD4 counts of all participants. If more than one clone represented a viral population that dominated at a particular time point then the entry efficiency of the clones was averaged. The association between relative *in vitro* Env efficiency and disease progression was analysed by Spearman r correlation test. Spearman r test is indicative of a negligible, weak, moderate, high and very high relationship based on previous reports (Hinkle et al., 2003) and a p value of < 0.05 was considered significant. GraphPad Prism 5.0 software (CA, USA) was used for statistical analysis.

2.5.5 Env Fitness determinants

The relationship between *in vivo* variant frequency and Env entry efficiency suggests that Env function influences viral fitness. We hypothesized that the better the virus is able to enter host cells, the faster they would replicate. Therefore, identifying determinants of Env fitness could provide novel targets for vaccine and drug design that if targeted, could lead to attenuation of viral replication and reduced risk of transmission. To identify Env fitness determinants that could potentially impact viral outgrowth, Env sequences were aligned and mismatches were compared across Env with varying entry efficiency. Results revealed putative sites that could play a role in influencing Env structure and function.

CAP84

At 3 mpi CAP84 was infected with two variants that differed primarily in gp41, suggesting that the gp120 of the second parent (virus B) was not detected. Virus A (84c1) had accumulated a number of mutations by 12 mpi, with 84c3 carrying a cluster of 9 amino acid changes within the signal peptide and putative CTL escape mutations within CAP84 HLA restricted epitopes. These changes seemingly reduced 84c3 entry efficiency significantly compared to 84c1 (Figure 2.9). However, after recombination introduced virus B regions within gp41 of virus A, the entry efficiency of virus AB (84c4) was significantly higher than 84c3 coincident with increased frequency (70 %). Virus A and virus AB at 12 mpi comprised very similar gp120, suggesting that gp41 might be influencing their PSV Env entry efficiency and thus driving *in vivo* viral outgrowth.

To investigate whether gp120 mutations lowered and gp41 increased PSV entry efficiency, we generated two chimeric Env clones, (84c1/c4 and 84c4/c1) by swapping the gp41 of virus A at 0 mpi for that of virus AB and *vice versa*. The chimera (84c1/c4) with gp41 of virus B was 3- to 4-fold significantly more efficient than PSV A at entering TZM-bl cells ($p = 0.005$). Despite high sequence similarity with 84c3, 84c4/c1 had lower entry efficiency than 84c1 and 84c4 although this difference did not reach significance (Figure 2.13 A and B). This suggested that recombination in gp41 may have resulted in a fitter virus but that mutations, present in gp120 of virus AB, attenuated Env function.

When gp120 sequence of virus A was compared to that of virus AB, a potential N-glycosylation site (PNG) at position 339 was identified in AB sequences that appeared in 19/20 sequences at 12 mpi. The introduction of N339 by SDM into the 84c1/c4 chimera with high entry efficiency decreased PSV entry 2.4-fold ($p < 0.05$), suggesting that an N-glycan at this position had a negative effect on Env function. We thus hypothesized that a PNG at position 339 of virus AB could counter the positive effect of recombination in gp41 on Env function. However, when N339 was deleted in 84c4, the entry efficiency did not increase as expected (Figure 2.13C), suggesting that the putative CTL escape mutations also had a deleterious effect on PSV entry efficiency.

A closer analysis of CAP84 sequences indicated that the introduction of N339 was concomitant with adjacent amino acid changes that flanked the CTL epitope recognized by HLA B*1503 (Appendix B). Mutations flanking CTL epitopes can prevent processing and presentation of peptides to CD8+ T cells facilitating immune escape (Milicic et al., 2005). Not only has the introduction of PNGs been linked to escape from neutralizing antibodies (Moore et al., 2012) and decreased Env fitness (Wang et al., 2013), it has also been linked to processing and presentation of CTL epitopes (Wood and Elliott, 1998). As CAP84 is HLA B*1503 positive and additional mutations appeared within and adjacent to other B*1503 epitopes (Appendix B), it is possible that the emergence of mutations and an N-glycan at N339 flanking the epitope VY9 contributed to immune escape at 12 mpi. However, these mutations might also have synergistically influenced Env structure and function within the context of 84c4 gp41. Therefore, deletion of N339 without altering the mutations flanking VY9 in 84c4 might not have resulted in a change in Env entry efficiency (Figure 2.13C). Collectively, this data suggests that the introduction of mutations in gp120 could have reduced Env fitness that was then rescued by recombination in gp41 and highlights the complex relationship between Env function and viral evolution.

CAP137

The outgrowth of virus AB at 12 mpi was associated with enhanced entry efficiency of PSV 137c10, a representative of the AB viral population. Therefore, in order to identify fitness determinants, AB sequences were compared to the master sequences representing virus A and virus B. Sequence analysis showed that virus AB had inherited regions/sites from both initial infecting viruses and could be sub-grouped into recombinants AB/A and AB/B due to the presence of additional gp41 sequence of virus B (Figure 2.6). Although, the difference in entry efficiency between recombinants AB/A and AB/B was not significant, with the concomitant 2-fold higher frequency of AB/B it seemed likely that this region in gp41 played an important role in virus AB/B PSV entry efficiency.

Further sequence analysis also identified a PNG at position 332 (according to HXB2 numbering) of virus B, AB/A and AB/B gp120. The PNG at N332 was identified as the epitope for broadly cross-neutralizing (BCN) antibodies that emerged during late stages of

infection (Moore et al., 2012). As the PNG elicits some of the most potent BCN antibodies and was inherited from virus B that had poor PSV entry efficiency, we wanted to know whether introduction of this PNG in AB virus was accompanied by a change in Env entry efficiency. When the PNG at N332 was deleted in 137c10, there was a significant 2.3-fold decrease in entry efficiency compared to wild-type ($p < 0.001$), suggesting that a PNG at position 332 might not only play a role in escape from neutralising antibodies but also enhance viral entry efficiency (Figure 2.14). However, noting that virus B had poor entry efficiency despite carrying the PNG, the affect that a PNG at 332 might have on Env entry efficiency is most likely dependent on the presence of other sequence changes. Therefore, it is impossible to predict whether immune responses that drive the loss or incorporation of PNGs will lead to less fit variants as fitness is likely highly dependent on the interplay between a number of structural factors.

CAP267

Based on the entry efficiency results, the evolution of virus B over time coincided with an increase in PSV entry efficiency. PSV representing virus B (267c6) at 12 mpi had 4.2-fold higher entry efficiency than PSV carrying virus B Env (267c2) at 0 mpi. These sequences differed by eighteen amino acids in gp120 and five amino acids in gp41. To determine whether gp120 or gp41 changes were responsible for the difference in entry efficiency, gp120 (excluding C5 region) was swapped between the two clones and entry efficiency was compared. The chimera having C5 and gp41 of virus B (267c6) at 12 mpi had 4.5-fold increase in entry efficiency, similar to that of PSV B (267c6) (Figure 2.15). This suggested that either the one a.a. in C5 and/or the five a.a. changes in the gp41 of virus B at 12 mpi could be responsible for the increase in entry efficiency of the B population over time.

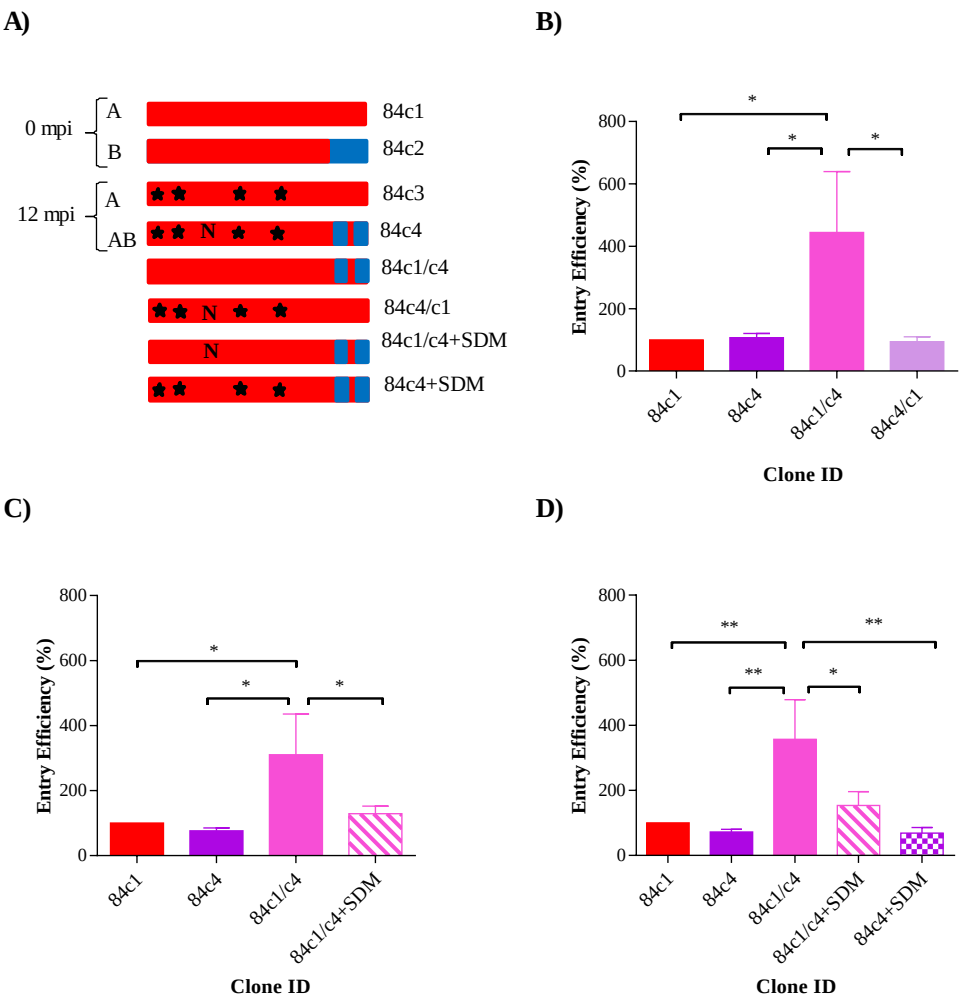


Figure 2.13. Entry efficiency of CAP84 Envelope chimeras and mutants. A) Schematic diagram of CAP84 clones, chimeric and mutant constructs made for CAP84. Red and blue bars indicate virus A and virus B sequences, respectively, stars indicate point mutations and N indicates the potential N-glycan site (PNG) N339. B) Two chimeras were constructed from 84c1 and 84c4 by swapping gp41 between the two clones. CAP84 virus A (84c1) at 0 mpi and virus AB (84c4) at 12 mpi are indicated in red and purple, respectively. The chimera 84c1/84c4 comprises of the gp120 region of 84c1 and the gp41 of 84c4 and is indicated by a pink bar. Chimera 84c4/c1 comprises of gp120 of 84c4 and gp41 of 84c1and is indicated by the pale purple bar. C) Site-directed mutagenesis (SDM) was used to introduce a PNG site at position 339 in the chimera 84c1/c4. The mutant is indicated by the pink striped bar. D) PNG site at position 339 was removed from 84c4 using SDM and is indicated by purple blocks. Entry efficiency is indicated as percentage relative to that of virus A (84c1). Data represents the average of three independent experiments. One-way ANOVA with Bonferroni correction for multiple comparisons was used for statistical analysis using GraphPad Prism 5.0. ($p < 0.05$ = *, $p < 0.01$ = **, $p < 0.001$ = ***, and $p < 0.0001$ = ****).

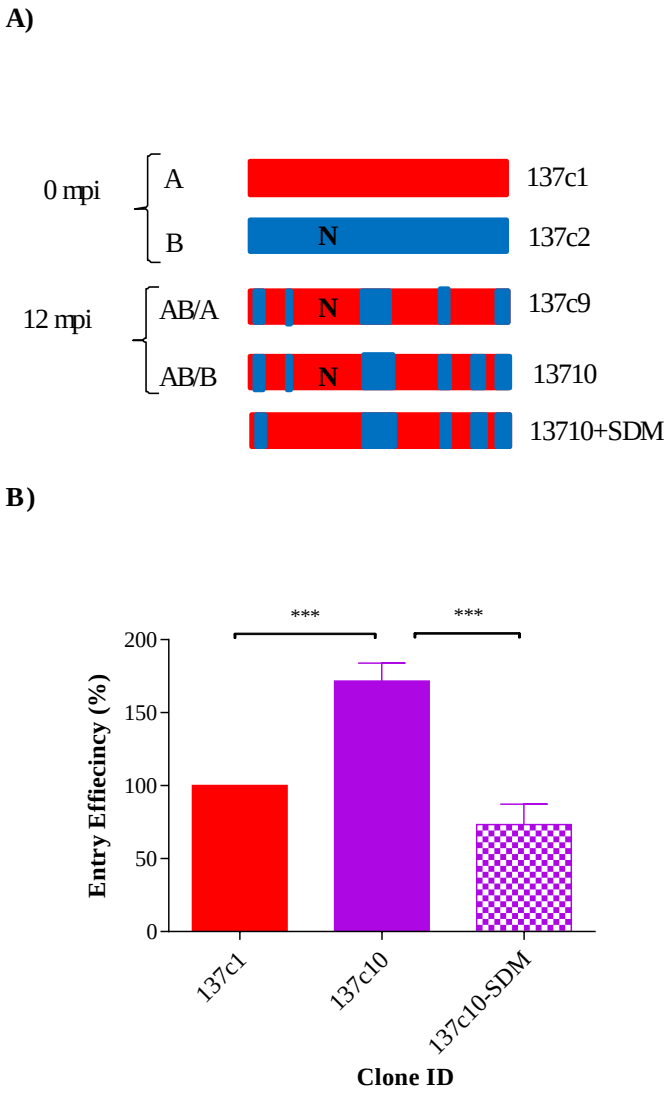


Figure 2.14. Identification of CAP137 Envelope determinants. A) Schematic diagram of CAP137 clones and mutant construct. Red and blue bars indicate sequences from virus A and virus B, respectively with recombined regions indicated. The N indicates the PNG, N332. B) A PNG was removed from virus AB/B (137c10) at position 332 (HXB2) and the entry efficiency was compared using the TZM-bl cell line. Virus A (137c1), virus AB (137c10), and the 137c10-sdm mutant are indicated by red, purple and purple blocks, respectively. Bars represent the average of three independent biological repeats with error bars representing standard deviation. One-way ANOVA with Bonferroni correction for multiple comparisons was used for statistical analysis using GraphPad Prism 5.0. ($p < 0.05$ = *, $p < 0.01$ = **, $p < 0.001$ = ***, and $p < 0.0001$ = ****).

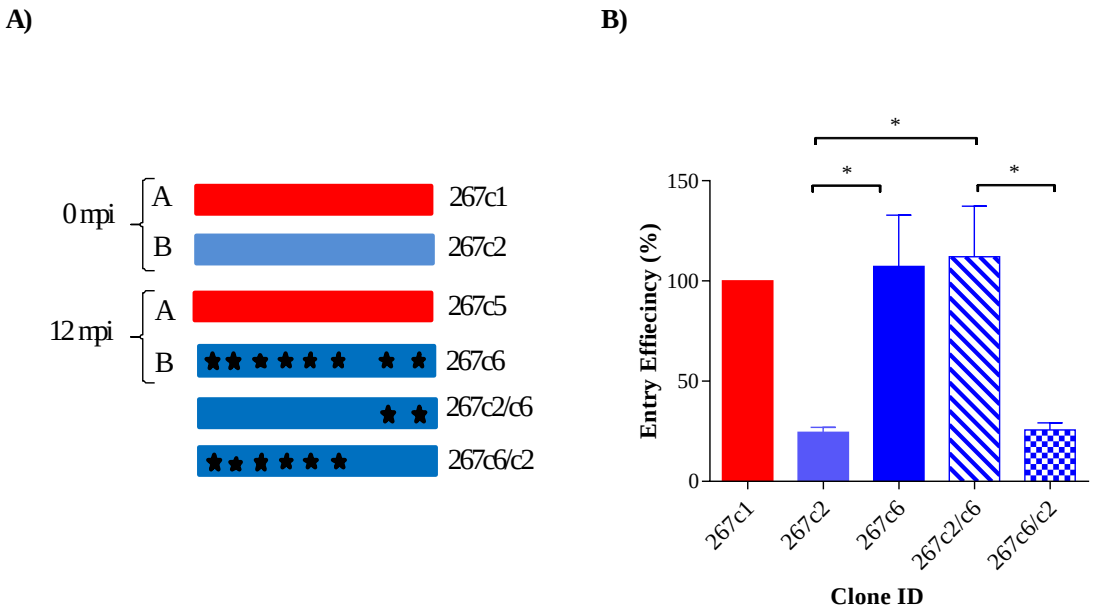


Figure 2.15. Entry efficiency of CAP267 chimeras. A) Schematic diagram of CAP267 clones and chimeras. Red and blue indicate the sequences of viruses A and B, respectively with stars indicating point mutations. B) Gp120 of virus B at 0 mpi (267c2) (blue bar) was swapped with that of virus B at 12 mpi (267c6) (blue bar). Two chimeras were generated using restriction enzymes and entry efficiency was compared to virus A at 0 mpi (267c1) indicated with a red bar. The first chimera (267c2/c6) comprises the gp120 of virus B at 0 mpi and gp41 of virus B at 12 mpi (blue striped bar). The second one (267c6/c2) comprises gp120 of virus B at 12 mpi and gp41 of virus B at 0 mpi (indicated by blue blocks). One-way ANOVA with Bonferroni correction for multiple comparisons was used for statistical analysis using Prism 5.0. Data represent the average of three independent biological repeats with standard deviation ($p < 0.05$ = *, $p < 0.01$ = **, $p < 0.001$ = ***, and $p < 0.0001$ = ****).

2.6 Discussion

Env is an important determinant of viral fitness and can impact disease progression (Quiñones-Mateu et al., 2000). We aimed to determine whether the in vivo outgrowth of variants infecting dual infected individuals was related to the *in vitro* Env entry efficiency of the dominant viruses at 12 mpi.

In order to track viral evolution and monitor fluctuations in viral populations, we first needed to identify the initial transmitted viruses. Identification of viral populations via sequencing is limited by sampling and as we analysed ~20 sequences per time point, we were most likely only detecting variants at a frequency greater than 15 % (Salazar-Gonzalez et al., 2008). Furthermore, although we had samples as early as 2-12 wpi, rapid recombination or transmission of recombinants, complicated the identification of full-length parent envs at the earliest time point of sampling (Wooley et al., 1997). To overcome this limitation, we determined the maximum DNA distance between the dominant variant at the earliest time point, virus A and all other variants. We hypothesised that the variant most dissimilar to virus A, would most closely resemble the second parent, virus B. Based on Highlighter plots; however, it was highly likely that the sequence we classified as virus B was a recombinant of A. Therefore, we suspected that recombination had occurred before the earliest time of sampling for all the dual-infected individuals. As we aimed to determine whether viral outgrowth was due to Env fitness, we hypothesised that viruses below detectable frequency would not significantly influence determining the relationship between in vivo viral outgrowth and Env entry efficiency. We identified two variants, virus A and virus B, most closely resembling the two phylogenetically distinct parents, tracked their relative frequency over time and determined whether Env function was driving viral outgrowth.

Viruses A and B from all participants differed with a DNA distance ranging between 3 – 12 %. In CAP84, env representing virus B was only detected at 3 mpi with a max DNA distance of 3 % while env classified later as a recombinant of viruses A and B was already

present at ~1 mpi. The inability to detect virus B variants at earlier time points suggested low in vivo frequency likely due to poor replication fitness.

Further recombination events between viruses A and B were confirmed using Highlighter Plot and RIP tools. The timing of the emergence of new recombinant viruses varied between participants. In two participants, recombinants emerged at 2 and 6 mpi (CAP37 and CAP137, respectively) whereas in CAP84, an apparent recombinant of viruses A and B was present at 1 mpi, before the detection of virus B and there were no detectable recombinants of virus A and virus B infecting CAP267. Despite differences in timing between participants, the *in vivo* frequency of recombinants increased over time and outgrew other variants at 12 mpi in CAP37, CAP84 and CAP137.

Based on previous definitions, the four dual-infected participants were classified as rapid (CAP37, CAP137, CAP267) and typical (CAP84) progressors based on their CD4+ T cell levels (Mlisana et al., 2014). In our study, the in vivo frequency of emergent recombinants in two rapid progressors at 12 mpi confirmed previous reports that showed an association between recombination and disease progression (Fang et al., 2004; Gordon et al., 2016; Nájera et al., 2002; Ramirez et al., 2008). Interestingly, no emergent recombinants were detected in the other rapid progressor (CAP267) although virus B, a putative recombinant, had the highest frequency at 12 mpi. However, in the typical progressor (CAP84), the early presence of recombinant variants and the in vivo outgrowth of emergent recombinants were not associated with disease progression. Although the sample size is small, and we cannot exclude the impact of limited sampling, these results could suggest that recombination is not always linked to disease progression (Liu et al., 1997).

Our results showed variation in PSV entry efficiency (ranged from 10 – 230 %) across participants. At early time-points, there was no consistent relationship between PSV entry efficiency and viral outgrowth. The entry efficiency of PSV representing variants at 12 mpi infecting CAP37 and CAP84 did not increase significantly compared to viruses at 0 mpi, suggesting that diversification had not selected for fitter Env. However, at 12 mpi 3/4 participants were infected with variants with high frequency (dominant) with significantly higher entry efficiency than other variants at the same timepoint, suggesting that Env fitness

was playing a role in viral outgrowth within later time points. The only exception was CAP37, where recombinant AB/B, with the highest entry efficiency at 12 mpi had similar frequency to recombinant AB/A at 12 mpi.

Whether outgrowth was due to gain of fitness of Env or whether recombination triggered the emergence of escape mutants that enhanced viral survival, required further analysis (Mostowy et al., 2011). As detailed analysis of immune escape fell outside the objectives of this study, we only looked at the emergence of polymorphisms within previously identified CTL epitopes recognised by the HLA alleles of the individuals. There seemed to be some evidence of toggling of polymorphisms during early infection with subsequent fixation of CTL escape mutations for CAP37, CAP84 and CAP137 by 12 mpi (Abrahams et al., 2013). It is interesting to speculate that the seemingly random changes in entry efficiency during early infection could be due to immune pressure selecting escape mutations that either favoured or lowered Env entry efficiency. However, based on CTL epitope analysis the recombinants at 12 mpi had not only escaped early immune responses but also gained fitness mutations favouring *in vivo* outgrowth although without further analysis no clear conclusions can be drawn. In our cohort, there was no increase in PNGs frequency over time, suggesting that bnAB might not play a role within the first year of infection as previously suggested (Abrahams et al., 2013).

CAP84 PSV representative of the emergent recombinant that outgrew at 12 mpi did not have enhanced entry efficiency compared to the parents despite apparent positive selection for changes in the signal peptide and gp41. When we generated gp120/gp41chimeras of virus A and virus AB, the gp41 of virus AB fused to the gp120 of virus A led to enhanced PSV entry, suggesting that the positive selection of virus B gp41 should have resulted in a virus with a fitter Env. Further sequence analysis identified a PNG at position 339 present in virus AB and absent in virus A. Changes in Env N-glycosylation in reaction to host immune responses has been linked to changes in Env fitness and we thus wanted to test whether the presence of a PNG at 339 affected CAP84 PSV entry efficiency. The introduction of N339 into the chimera, decreased PSV entry efficiency, suggesting that the lack of increased entry efficiency of CAP84 emergent recombinants could be due to a fitness cost associated with N339 (Liu et al., 2007). However, deletion of the PNG in virus AB at 12 mpi (84c4) did not

yield the expected increase, suggesting that the impact of mutations on Env fitness is variant specific.

The appearance of the PNG at position 339 was also co-incident with mutations flanking the VY9 epitope recognized by HLA B*1503, an allele frequents in subtype C sequences and associated with slow disease progression in subtype B but not subtype C infected individuals (Frahm et al., 2006). It is thus possible that the concomitant emergence of the PNG at N339 and potential CTL escape mutations was required for changes in entry efficiency. This has ramification for vaccine design as PNGs that help escape from neutralizing antibody recognition and potentially CTL escape could increase the fitness of one variant in one host but attenuate Env in another. This is especially noteworthy given that N-glycans are the target for potent neutralizing antibodies (Lavine et al., 2012). We propose that recombination within gp41 increased the fitness of emergent recombinants but that due to immune pressure a PNG site was introduced at position 339 which then attenuated viral replication. The lack of overall enhanced entry efficiency over the first 12 months of infection could be responsible for the typical disease progression of this participant (Ball et al., 2003; Quiñones-Mateu et al., 2000).

Similarly, for CAP37 the entry efficiency of PSV representing variants at 12 mpi was similar to that at 0 mpi, suggesting no overall increase in viral fitness. However, the entry efficiency of viruses A and B declined over time and were supplanted by the emergent recombinants AB/A and AB/B, most likely due to escape from immune responses (Barouch and Letvin, 2002; Costa et al., 2004; Moore et al., 2002). AB/A had the highest entry efficiency at 12 mpi, suggesting that recombination had selected for a fitter virus at 12 mpi but escape from immune responses compromised Env entry efficiency. There was evidence of putative CTL escape mutations within epitopes recognized by HLB*5301.

In the other two participants, CAP137 and CAP267, PSV entry efficiency increased over time coinciding with in vivo outgrowth of variants. Env entry efficiency was reported to be linked with disease progression (Gordon et al., 2016; Lassen et al., 2009; Quiñones-Mateu et al., 2000; Trkola et al., 2003). Here, we showed that PSV entry efficiency correlated significantly with faster CD4+ T cell decline ($p = 0.046$) and therefore, Env entry efficiency

could be a driving factor in the disease progression of CAP137 and CAP267. A similar relationship between PSV entry efficiency and VL was not observed but CD4+ T cell decline can be an independent marker of disease progression and in vivo variant frequency (Claiborne et al., 2015). Furthermore, when the entry efficiency of the dominant variant at each time point was correlated to CD4 counts, the association was strengthened, supporting the relationship between Env fitness and disease progression.

Interestingly, gp41 was identified as a potential fitness determinant of variants infecting CAP137, CAP267 and CAP84, at 12 mpi, suggesting that gp41 could play a very important role in Env entry efficiency by determining virus-to-cell membrane fusion (Checkley et al., 2011; Wilen et al., 2012). However, Env fitness of CAP137 AB/B was also influenced by a PNG at position 332. It has been shown that broadly neutralising antibodies recognised an N-glycan at position 334 and that escape involved the shifting of the PNG from position 334 to 332. However, shifting of the PNG to position 332 resulted in broadly cross-neutralising (BCN) antibodies so it was unclear how a PNG at 332 benefitted the virus (Moore et al., 2012). When we deleted N332, the enhanced entry efficiency of the emergent AB/B recombinant dropped significantly, suggesting that a PNG at position 332 might not only play a role in immune responses but also Env fitness which is consistent with a previous finding (Lavine et al., 2012) where N332S also reduced virus entry of HIV-1JR-FL. Thus, this could explain the benefit of a shift of PNG from 334 to 332 although whether a PNG at position 334 also influences Env function still needs to be determined.

Conclusion

In this study, we showed that recombinants were dominant at 12 mpi with better entry efficiency compared to other viruses at the same time point. The increased entry seemed to be associated with recombination within gp41 for 3/4 individuals. However, for CAP84 and CAP137 changes in gp41 were linked to changes in PNGs at positions 339 and 332 of the immunogenic C3 region (Moore et al., 2008), suggesting that interplay between immune escape and changes in Env function could determine the outgrowth of fitter variants. As Env entry efficiency could play a very important role in CD4 T cell loss, vaccines that target Env

N-glycans could select for the loss or gain of PNGs resulting in the emergence of a more pathogenic variant that increases disease progression.

CHAPTER 3.

Longitudinal, Phenotypic Characterisation of Hiv-1 Subtype C Envelope Isolated from Dual Infected Individuals

3.1 Introduction

In this study, recombination events, which mapped mainly to gp41 in association with N-glycan site changes in the C3 region, were associated with *in vivo* outgrowth of viruses (Chapter 2). Outgrowth of recombinants was associated with enhanced Env entry efficiency in 2/4 dual infected individuals and mutagenesis showed that changes within C3 and gp41of ¾ individuals conferred phenotypic advantages that enhanced Env entry efficiency. We hypothesised that changes in Env phenotype mediated by C3 and gp41 could translate into high replicative fitness which might increase disease progression. Env phenotypic changes that have been associated with HIV-1 pathogenicity include: (i) switching of R5-tropism to X4 tropism (Gorry et al., 2014; Gorry and Ancuta, 2011), ii) enriched macrophage (M)-tropism and high replicative capacity in monocyte derived macrophages (MDM) (Blaak et al., 2000; Li et al., 1999), (iii) increased ability to utilize low levels of CD4 (Gorry et al., 2014) and CCR5 (Gray et al., 2005; Tuttle et al., 2002) with greater susceptibility to inhibition by CCR5 antagonists (Koning et al., 2003; Lobritz et al., 2007; Olivieri et al., 2007), and (iv) enhanced Env entry efficiency and viral fitness (Marozsan et al., 2005; Rangel et al., 2003). We therefore wanted to determine how recombination in gp41 and/or the presence of PNGs influences Env entry efficiency and thus viral fitness. As gp41 seemed to be the most common fitness determinant across participants, we focused on phenotypes related to gp41 function. However, as PNGs at sites 332 and 339 were adjacent to V3, involved in R5 tropism and binding to CCR5 (Bagnarelli et al., 2003; Hwang et al., 1991), we also investigated differences in co-receptor and cellular tropism.

Following Env interaction with CD4 and CCR5, a series of conformational changes in gp41 occur that result in the fusion of the viral and cell membranes (Chan et al., 1997; Pancera et al., 2010, p. 41). Previous studies showed that differences in fusion capacity of HIV-1 variants impact virus entry (Lobritz et al., 2007; Marozsan et al., 2005), and the sensitivity to fusion inhibitor (T-20) (Lobritz et al., 2007; Reeves et al., 2002). These differences in fusion capacity were mapped to variable loops (V1-V3) in gp120 (Cavrois et al., 2014; Etemad et al., 2009; Lobritz et al., 2007) and the T-20 binding site on the HR1 domain of gp41 (Reeves et al., 2005; Wei et al., 2002). Moreover, previous studies also revealed an association between fusion capacity and post-translational processing and incorporation of cleaved Env into virions (Blay et al., 2007; Cavrois et al., 2014) resulting in enhanced Env entry efficiency (Bachrach et al., 2005). Env incorporation into virions was mapped to gp41 (Blay et al., 2007; Da Silva et al., 2013; Murakami and Freed, 2000), emphasising the important role of gp41 in Env fitness.

Therefore, we compared Env phenotypes between clones that varied in PSV entry efficiency to determine one or more Env properties that could provide variants with a functional advantage over another in dual infected individuals.

3.2 Research Aim and Objectives

Aim

To determine the phenotype of Env driving changes in pseudovirus entry efficiency over 12 months of dual infection.

Objectives

Objective 1: Compare changes in Env tropism, fusion capacity, expression and incorporation into virions over the course of infection.

Objective 2: Identify Env phenotype(s) responsible for changes in pseudovirus entry efficiency.

Objective 3: Determine the impact of fitness determinants on changes in Env phenotypes.

3.3 Materials and Methods

3.3.1 Samples used in this study

A subset of clones isolated from CAP84, CAP137 and CAP267 with changes in entry efficiency between enrolment and 12 mpi were selected.

3.3.2 Coreceptor phenotype

To determine Env viral tropism in our cohort, *envs* were co-transfected with the luciferase-encoding pNL4-3.Luc.R^-E^- backbone (NIH AIDS Reagent Program, Division of AIDS, NIAID, NIH: pNL4-3.Luc.R^-E^- from Dr. Nathaniel Landau, USA). U87 cell lines expressing CD4 receptor and either CCR5 or CXCR4 were infected with PSV with a titre equivalent to 100 ng/ml of p24 and luciferase activity was measured after 48 hours. Viruses that could infect the U87CD4+CCR5+ cells (NIH AIDS Reagent Program, Division of AIDS, NIAID, NIH: U87CD4+CCR5+ cells from Dr. HongKui Deng and Dr. Dan R. Littman) were classified as R5-tropic, and viruses that infected U87CD4+CXCR4+ (NIH AIDS Reagent Program, Division of AIDS, NIAID, NIH: U87 CD4+CXCR4+ from Dr. HongKui Deng and Dr. Dan R. Littman) were classified as X4-tropic. In order to control for R5 and X4 tropism, we included the Env clones, QHO (Li et al., 2005) to control for R5 tropism and RPI (Cilliers et al., 2005) for X4 tropism. Transfection with the pNL4-3.Luc.R^-E^- backbone and medium only was used as negative controls. Coreceptor usage was also confirmed using the bioinformatics tool, Position-Specific Scoring Matrices (PSSM) (https://indra.mullins.microbiol.washington.edu/webpssm/) (Jensen et al., 2003), that predict R5 and X4 usage based on the V3 positive charge.

3.3.3 Cellular tropism

HEK293T- Affinofile cells, donated by Benhur Lee (UCLA) are dual inducible so that the addition of minocycline (Sigma-Aldrich, MO, USA) and ponesterone (PonA) (Sigma-Aldrich, MO, USA) induces CD4 and CCR5 levels in a dose dependent manner. These cells were used to measure whether Envs were macrophage-tropic (M-tropic) or T lymphocyte tropic (T-tropic) based on the ability of PSVs to enter cells expressing low

CD4 and high CCR5 levels resembling macrophages or high CD4 and low CCR5 levels similar to T cells (Table 3.1) (Johnston et al., 2009). Briefly, Affinofile cells (10^4 cells per well) were seeded in 6 well plates, CD4 expression was induced with a concentration range of minocycline (0-5 ng/ml) and CCR5 expression was induced with a range of PonA (0-4 μM). After inducing for 24 hours, cells were collected and stained with either phycoerythrin (PE)-labelled anti-human CD4 or PE-labelled anti-human CCR5 (BioLegend, US) for 30 minutes at room temperature. Cells were fixed with paraformaldehyde and CD4 and CCR5 density was measured using flow cytometry. Quantibrite beads (BD Biosciences, USA) were used for CD4 and CCR5 quantification according to the manufacturer's instructions. The number of antibody sites/cell (ABS/cell) was equivalent to molecules per cell (Joseph et al., 2014). Data acquisition and flow cytometry analysis was done using Cellquest and Flowjo software, respectively.

Cellular tropism was measured by seeding Affinofile cells (10^4 cells/well) in a 96-well plate and incubated at 37 °C for 24 hrs. Three different combinations of antibiotics were used to induce different densities of CD4 and CCR5 (Table 3.1). T-lymphocyte tropism was determined by inducing cells with the required concentration of minocycline and PonA to express high CD4 (~70,000 molecules/cell) and low CCR5 (~2,000 molecules/cell), respectively. Affinofile cells were also induced with the required concentration of minocycline and PonA to express low CD4 density (~2,500 molecules/cell) and high CCR5 (~31,000 molecules/cell) to resemble macrophage levels (Ping et al., 2013). Affinofile cells were also induced to express high density of CD4 (~110,000 molecules/cell and a range of levels of CCR5 from (2,000 – 31,000 molecule/cell). The cells were incubated for 24 hours before infection with PSV equivalent to 100 ng/ml p24 and entry efficiency was determined after 48 hrs using a Briteglo kit (Promega, USA) according to the manufacturer's instructions and luminescence was measured on a luminometer (Turner Biosystems® Modulus Microplate). JR-CSF (obtained from Lynn Morris, NICD, South Africa) was included as a representative of T-tropic variants. Cells-only (mock infection) was used as background.

Table 3.1. Density of CD4 and CCR5 induced by Minocycline and Ponesterone A

Cell Type	CD4			CCR5			[a] Physiological receptor concentration (molecule/cell)	
	Minocycline ng/ml	Molecules/ cell	[b] Published	Ponasterone A µM	Molecules/ cell	[b] Published	CD4	CCR5
Affinofile T-lymphocyte-like: $CD4^{High}/CCR5^{low}$	2.5	$7x10^4$	$7x10^4$– $1.1x10^5$	0.06	$2x10^3$	$1x10^3 – 2.7x10^3$	$6.5x10^4$– $4.5x10^5$	10^3 – $1.2x10^4$
Affinofile Macrophage-like: $CD4^{Low}CR5^{high}$	0.6	$2.6x10^3$	$2x10^3$	0.01	$3.1x10^4$	$3.0x10^4$	~10^4	$8x10^3$ – $8.4x10^4$
TZM-bl	N/A	$2.5x10^5$	$6.5x10^4$	N/A	$1.1x10^4$	$2x10^3 – 1x10^5$	N/A	N/A
U87 CD4+/CCR5+	N/A	$2.0x10^5$	NF	N/A	$6.0x10^3$	NF	N/A	N/A

[a] Refers to concentration of CD4 and CCR5 on T-lymphocytes and Macrophages (Johnston et al., 2009; Lee et al., 1999; Platt et al., 1998)
b (Johnston et al., 2009; Ping et al., 2013)
NF = not found; N/A: not applicable

3.3.4 Env sensitivity to entry inhibitor T-20

PSVs generated as previously described in section 2.3.3, were pre-incubated with five-fold serial dilution of T-20 (0.00032 – 5 µg/ml) (NIH AIDS Reagent Program (ARP), Division of AIDS, NIAID, NIH: T-20, (Enfuvirtide)) in DMEM for 1 hr at 37 °C prior to infection of TZM-bl cells (10^4 cells per well). Pseudotyped viruses were produced by co-transfecting the *env* clones with pSG3Δenv as previously described in Chapter 2. Luciferase activity was measured after 48 hours. RLU versus drug concentration were plotted to determine the 50 % inhibitory concentration (IC50) values for each PSV using the GraphPad Prism software 5.0 (CA, USA).

3.3.5 Expression and incorporation of Envelope into pseudovirus

To evaluate Env expression by HEK 293T cells and Env density on the surface of pseudoviral particles, cells expressing Env and pseudovirus were prepared, lysed and resolved on SDS PAGE followed by quantitative analysis using immunoblotting as described below.

3.3.5.1 Pseudovirus preparation

PSVs were generated as previously described in (Section 2.3.4.2, Chapter 2) with the following exceptions: $2X10^6$ cells/ml HEK 293T cells were plated in 10 cm tissue culture dishes (Corning, Sigma-Aldrich, MO, USA) and incubated at 37 °C with 5 % CO_2 for 24 hr after which cells were co-transfected with 5 µg of Env and 10 µg of the subtype B HIV-1 backbone, pSG3Δenv. Forty-eight hours later, pseudovirus supernatant was harvested and filtered through 0.2 µm filter and the cells were washed with 1X Phosphate-buffered saline (PBS). PSV was pelleted by centrifugation at 26000 X g for 2 hrs using a 20 % glycerol cushion at 4 °C in a RW55 rotor (Beckman Coulter, GER). The pelleted PSV was re-suspended in 50 µl PBS and stored at -80°C for further analysis.

Cells were lysed in 250 µl RIPA lysis buffer (Appendix C) for 5 - 10 min at 4 °C before centrifugation at 14000 rpm for 2 min. The supernatants were stored at -20 °C. Protein concentration was determined using a Bradford assay (Bio-Rad, USA) according to the

manufacturer's instructions. Bovine serum albumin (BSA) was used to construct a standard curve.

3.3.5.2 Immunoblotting

An equal volume of pelleted PSV and 80 µg of protein lysate was resolved using 8 % and 10 % SDS PAGE, respectively. Following electrophoresis, the proteins were transferred to polyvinylidene difluoride (PVDF) membrane (Immun-Blot®, Bio-Rad, USA) using the wet transfer system (Bio-Rad) and blocked overnight with 5 % blocking buffer (Appendix C) at 4 °C.

For the detection of pseudovirus-associated Env, the membrane was cut into two and the upper part was probed with sheep anti-gp120 (Table 3.2) for 1 hr and the lower section incubated with rabbit anti-p24 (Table 3.2) for 2 hr at room temperature. For detection of cell-associated Env, the lower section of the membrane was probed with mouse anti-actin antibody (Sigma-Aldrich, MO, USA) for 1 hour at room temperature. The membranes were washed three times with TBS-T, each for 15 min intervals, before incubation with secondary antibodies for 1 hr at room temperature (Table 3.2). Membranes were washed three times with TBS-T followed by one extra wash with TBS only as tween was found to interact with the detection reagent, LumiGlo chemiluminescent substrate (KPL) before visualisation using autoradiography film (Santa Cruz Biotechnology®, USA). Recombinant gp120 (NIH AIDS Reagent Program (ARP), Division of AIDS, NIAID, NIH: HIV-1 IIIB gp120 Recombinant Protein from ImmunoDX, LLC) was used as a positive control whereas mock-transfected cells and pseudovirus lacking Env were used as negative controls. The relative integrated intensity of Env, p24 and β-actin were determined using GelAnalyzer 2010 program and background was user defined.

3.3.6 Data analysis

Data represents the average of two independent Western blots. Based on the relative intensity of each band given by GelAnalyzer 2010, the expression and incorporation of Env into PSVs were calculated as described in (Blay et al., 2007; Provine et al., 2009) using the following formulae:

Cell-associated Env

(The band intensity of gp120 + band intensity of gp160) / (band intensity of β-actin).

Env expression and cleavage into gp120 and gp41

(Band intensity of gp120) / (band intensity of gp120 + band intensity of gp160)

Env incorporation into PSV

(The band intensity of gp120 + band intensity of gp160) / (band intensity of p24).

Table 3.2. Antibodies used in Western blot

Antibody Name	Protein target	Antibody type	Dilution	Source
Sheep anti-gp120 Mouse	Env gp160 and gp120	Primary	1:10 000	[a]ARP
Rabbit anti-p24 ARP432	Gag/ p24	Primary	1:10 000	[b]ARP
Mouse Anti-actin	β-Actin	Primary	1:4 000	Santa Cruz Biotechnology®
Goat anti-sheep IgG	[c] Sheep IgG antibody	Secondary	1:4 000	Santa Cruz Biotechnology®
Donkey anti-rabbit IgG	[c] Rabbit IgG antibody	Secondary	1:10000	Santa Cruz Biotechnology®
Goat anti-mouse IgG	[c] Mouse IgG antibody	Secondary	1:4 000	Santa Cruz Biotechnology®

[a] The antibody was obtained through the NIH AIDS Reagent Program (ARP), Division of AIDS, NIAID, NIH: Antiserum to HIV-1 gp120 from Dr. Michael Phelan (ARP)
[b] Antiserum to HIV-1 p24, ARP432 (ARP432, NIBSC Centralised Facility for AIDS Reagents, Hertfordshire, UK)
[c] Horseradish Peroxidase (HRP) conjugate.

3.3.7 Statistical analysis

One-way ANOVA test was used to compare the mean values of PSV Affinofile entry and fusion capacity between Env clones within each participant with Bonferroni test for multiple comparisons. In addition, Kruskal-Wallis test with Dunn's multiple comparisons test was

performed to compare mean Env expression and incorporation within each participant. Spearman correlation test was performed to determine the relationship between entry efficiency and all other phenotypes properties of Envs using GraphPad Prism 5.0 software (CA, USA). Correlation co-efficient (r) was determined as indicative of a negligible, weak, moderate, high and very high relationships based on previous reports (Hinkle et al., 2003) p value of < 0.05 was considered significant.

3.4 Results

3.4.1 Coreceptor phenotype

Coreceptor switch from CCR5 to CXCR4 was reported to be associated with an increase in viral load, CD4+ T cell depletion, Env fitness and therefore, disease progression (Connor et al., 1997; Gorry and Ancuta, 2011). We selected clones at 0 mpi and at 12 mpi from each participant to determine whether the coreceptor switch played a role in changes in Env entry efficiency of variants isolated from dual infected individuals. As the entry efficiency of PSVs at 3 and 6 mpi did not vary significantly from 0 mpi these clones were excluded from the analysis. All PSVs were only able to infect U87 cells expressing CD4 and CCR5 and not CD4 and CXCR4 (Figure 3.1). This was confirmed by the positive charge of V3 sequences using the PSSM tool (Data not shown). Variation between CAP37 and CAP137 clones infecting U87 cells expressing CD4 and CXCR4 could be due to experimental error as background values ranged from 50-300 RLU, suggesting inherent intra- and inter-assay variability (data not shown) similar to previous report (Whitcomb et al., 2007). Therefore, there was no change in coreceptor tropism over time, suggesting that co-receptor tropism did not contribute to changes in PSV entry efficiency.

3.4.2 Differences in cellular tropism

T-tropic viruses are able to infect cells with high CD4 and low CCR5 numbers while M-tropic have the ability to infect cells with low CD4 and high CCR5 levels (Gorry et al., 2002; Ping et al., 2013; Schnell et al., 2011). Recent studies have reported that transmitted founders (TFs) are T-tropic (Ping et al., 2013) but R5 viruses might evolve to M-tropism and infect host cells with low levels of CD4 (Beauparlant et al., 2017; Chikere et al., 2013). This will expand the range of HIV-permissible host cells, facilitating replication which could culminate in increased viral load and disease progression (Blaak et al., 2000; Chikere et al., 2014; Joseph et al., 2014). Therefore, to determine whether variants infecting the four dual

infected individuals were able to infect cells with low levels of CD4, PSV Env co-dependency on CD4 and CCR5 levels was measured by inducing 293 Affinofile cells to express different levels of CD4 and CCR5 simultaneously and independently (Johnston et al., 2009).

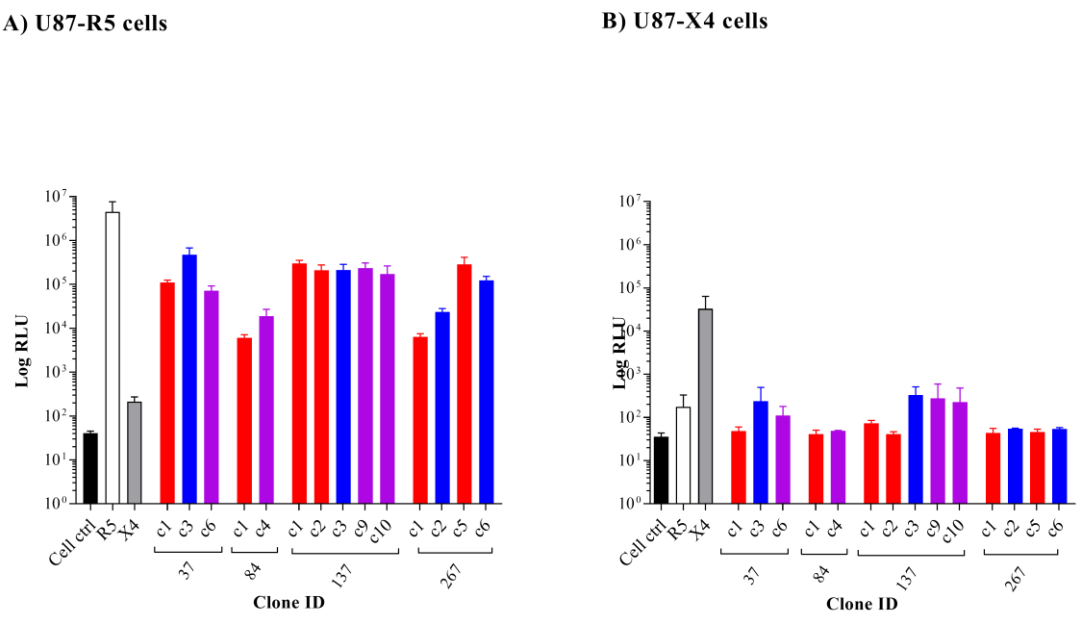


Figure 3.1. Determine coreceptor tropism over time. The coreceptor usage of pseudovirus (PSV) representing virus A, B and AB infecting CAP37, CAP84, CAP137, and CAP267 was compared over time (0, and 12 mpi except 137c3 at 2 mpi) in A) U87-R5 cell line and B) U87-X4 cell line. The coreceptor usage of PSV A, B and AB are indicated in red, blue and purple, respectively. Coreceptor usage was compared between PSVs normalised to 100 ng/ml p24 and displayed as log of relative luminescent units (RLU). Results represent the mean of two independent biological repeats with error bars indicating standard deviation. QHO (Li et al., 2005) (white bar), RPI (Cilliers et al., 2005) (gray bar) are laboratory strains that were used to control for R5, and X4 tropism, respectively. Cell control (Cell ctrl) indicated in black represents background of cells without virus (Mock).

Previous studies induced Affinofile cells to express a range of CD4 levels from low (~2,000 molecules/cell) to high (~70,000–110,000 molecules/cell) to mimic macrophages or activated T-lymphocytes, respectively, (Johnston et al., 2009; Ping et al., 2013). Affinofile

cells were also induced to express a range of CCR5 from low (2,691 molecules/cell) (Joseph et al., 2014) to high (~30,000 molecules/cell) (Ping et al., 2013) levels to mimic T-lymphocytes and macrophages, respectively (Table 3.1). In this study, Affinofile cells were induced to express (i) high levels of CD4 (~70,000 molecules/cell) and low levels of CCR5 (~2,000 molecules/cell) to mimic T-lymphocyte ($CD4^{high}/CCR5^{low}$), and (ii) low levels of CD4 (~2,600 molecules/cell) and high levels of CCR5 (~31,000 molecules/cell) similar to macrophages ($CD4^{low}/CCR5^{high}$) (Table 3.1).

To validate our Affinofile system, we used JR-CSF, as a T-tropic virus. JR-CSF infected Affinofile cells expressing high CD4 and low level of CCR5 ($CD4^{high}/CCR5^{low}$) efficiently and could not infect cells expressing low CD4 and high CCR5 ($CD4^{low}/CCR5^{high}$), suggesting that our Affinofile system was representative of T- cells and macrophages.

All PSVs from the four participants could enter and infect cells expressing high levels of CD4 despite low levels of CCR5 ($CD4^{high}/CCR5^{low}$) (Figure 3.2A-D). However, when the CD4 level was lowered to 2,600 molecules/cell and CCR5 levels were high ($CD4^{low}/CCR5^{high}$) (31,000 molecules/cell), RLU was not significantly higher than background. This data indicated that all Env clones required high levels of CD4 to enter cells and thus, were all T-tropic (Ping et al., 2013).

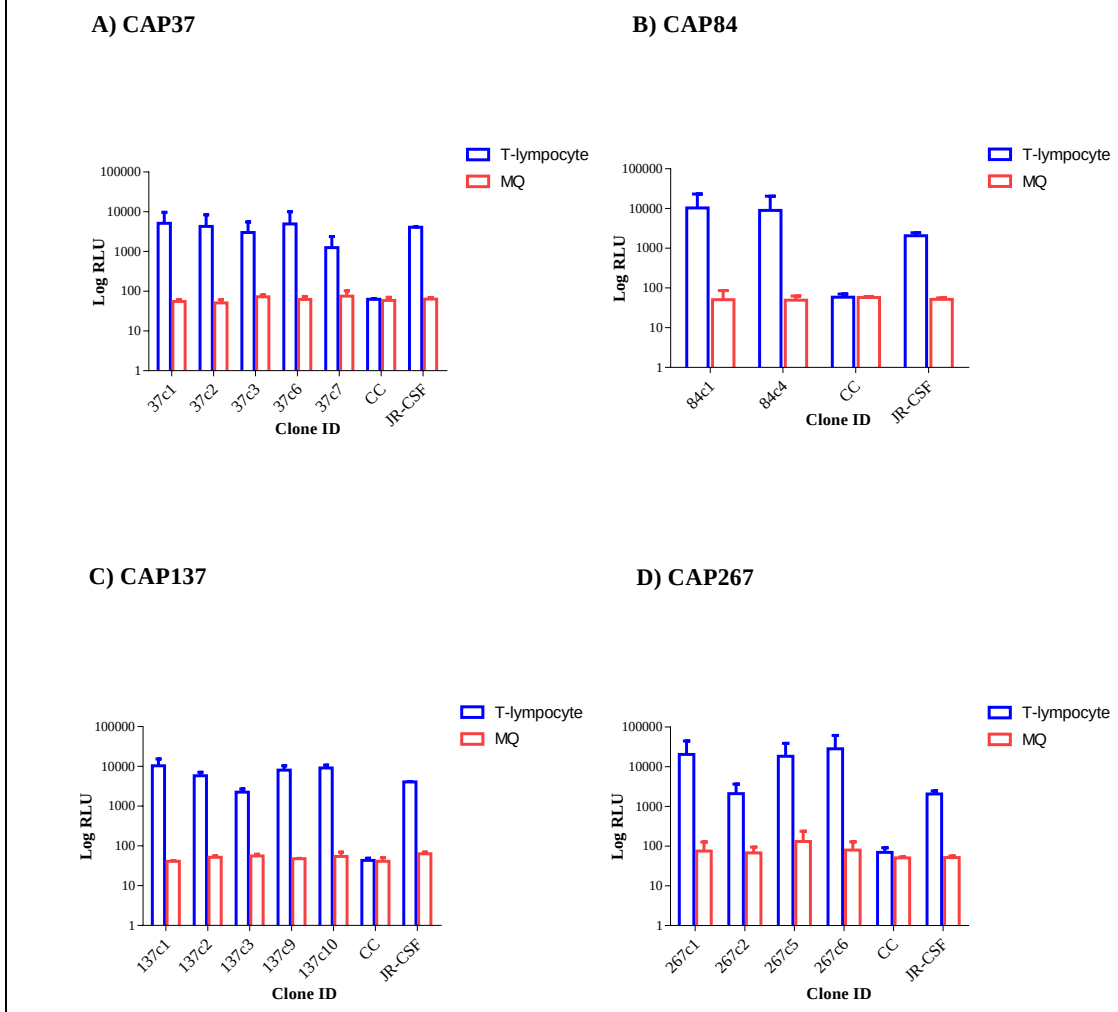


Figure 3.2. Cellular tropism using dual inducible 293 Affinofile cells. Pseudovirus (PSV) representing virus A, B and AB of A) CAP37, B) CAP84, C) CAP137, and D) CAP267 was compared over time (0, and 12 mpi except 137c3 at 2 mpi) using dual inducible Affinofile cells lines induced to express receptor levels that mimic T-lymphocytes (CD4 = 70,000 molecules/cell and CCR5 = 2000 molecule/cell) and macrophages (CD4 = 2,600 molecules/cell and CCR5 = 31,000 molecules/cell). High CD4 and Low CCR5 levels (T-tropism) are indicated by blue bars, Low CD4 and High CCR5 levels (M-tropism) are indicated in red. The entry efficiency was compared between PSVs normalised to 100 ng/ml p24 and displayed as log relative luminescent units (RLU). Results represent two independent biological repeats with error bars indicating standard deviation. JRCSF (T-tropic) and negative mock infection control (CC) are represented.

3.4.3 Validity of TZM-bl cells to compare Pseudovirion entry efficiency

In this study, TZM-bl cells were used to represent viral entry *in vivo* and U87-CCR5+ cells (Figure 3.3) and Affinofile cells (Figure 3.4) were utilised to characterise different PSV phenotypes. The TZM-bl assay required pSG3delenv HIV-backbone whereas the other two assays utilised pNL4-3.Luc.R^{-}E^{-}. As cell lines might not be physiologically relevant and viral-viral protein interactions might be important we compared the entry of clones in all three cell lines (Table 3.3). CD4 and CCR5 utilisation efficiencies have been shown to influence virus entry and infectivity (Chikere et al., 2013; Gray et al., 2005; Lee et al., 1999; Platt et al., 1998) but as CD4 levels were high across all cell lines, we did not consider that a limiting factor. We therefore, hypothesized that changes in the relative ranking of the clones between cell lines would indicate whether PSV entry efficiency differed due to HIV backbone and/or utilization of CCR5.

Table 3.3. Ranking of PSV entry efficiency according to cell type

ID	[a]TZM-bl ($CD4^{high}/CCR5^{higher}$)	[a,b]U87-CD4+CCR5+ ($CD4^{high}/CCR5^{high}$)	[b]Affinofile ($CD4^{high}/CCR5^{low}$)	*Sensitive PSV
CAP37	c3>c1≈c6	c3>c1≈c6	c1≈c6>c3	c3
CAP84	c1≈c4	c4>c1	c1≈c4	c4
CAP137	c10>c9≈c1>c2≈c3	c10≈c9≈c1≈c2≈c3	c10≈c9≈c1>c2>c3	c2, c3,c10
CAP267	c6>c5≈c1>c2	c5>c6>c2>c1	c6>c5≈c1>c2	c1, c5

[a] High CCR5 levels
b Same HIV backbone: pNL4-3.Luc.R–E–
*Identity of the Env clone(s) that differed in ranking in 1/3 cell lines

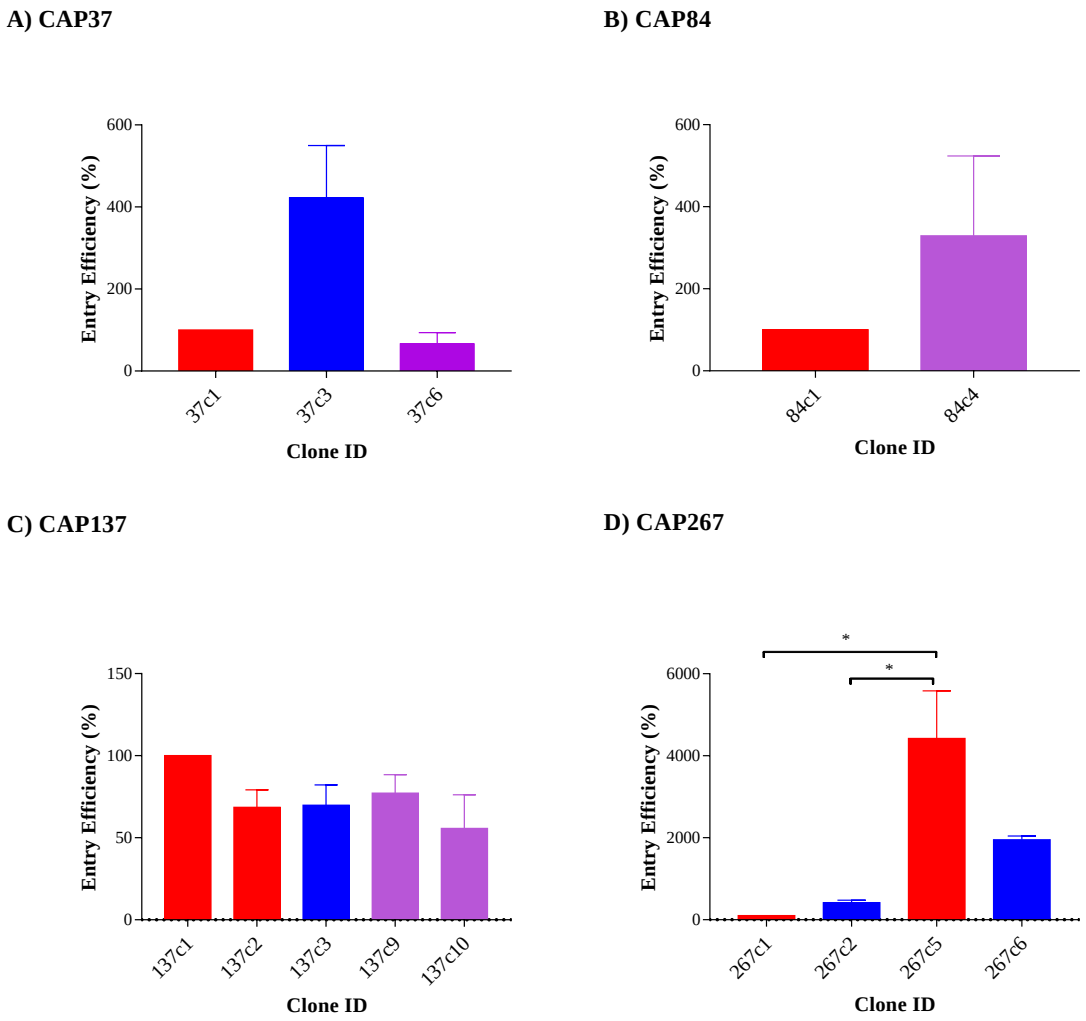


Figure 3.3. PSV entry efficiency in U87 CD4+CCR5+ cell line. The entry efficiency of pseudovirus (PSV) representing virus A, B and AB infecting A) CAP37, B) CAP84, C) CAP137, and D) CAP267 was compared over time (0 and 12 mpi and 2 mpi for 137c3) using U87 CD4+CCR5+ cell line. The entry efficiency of PSV A, B and AB are indicated in red, blue and purple, respectively. PSV entry efficiency was normalised to 100 ng/ml p24 and displayed as a percentage (%) relative luminescent units (RLU) to PSV A at 0 mpi. Results represent two independent biological repeats with error bars indicating standard deviation. Significant difference is indicated by stars. One-way ANOVA with Bonferroni correction for multiple comparisons was used for statistical analysis ($p < 0.05$ = *, $p < 0.01$ = **, $p < 0.001$ = ***, and $p < 0.0001$ = ****). All Env clones were functional as PSV entry efficiency was higher than the background control when no PSV was added to the cells. Infection with PSV generated by transfecting HIV-1 pNL4.3luc+E-R- backbone with an empty vector was used as negative control for infectivity and subtracted from all RLU values.

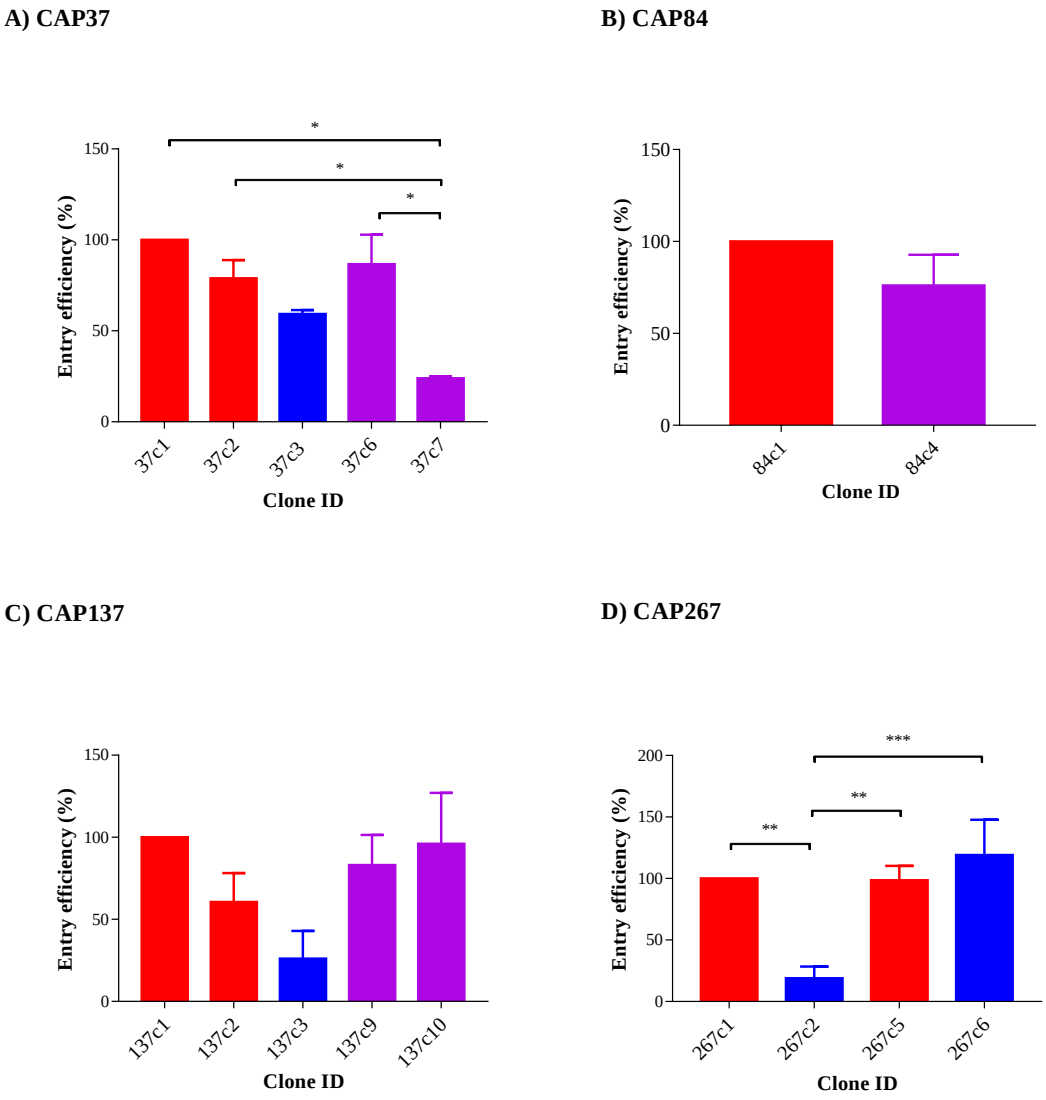


Figure 3.4. PSV Entry Efficiency in Affinofile cells expressing CD4^{high}/CCR5^{low}. Entry efficiency of pseudovirus (PSV) representing virus A, B and AB infecting A) CAP37, B) CAP84, C) CAP137, and D) CAP267 was compared over time (0, 2, and 12 mpi) using the Affinofile cell line expressing CD4^{high}/CCR5^{low}. The entry efficiency of PSV A, B and AB are indicated in red, blue and purple, respectively. The entry efficiency was compared between pseudoviruses normalised to 100 ng/ml p24 and displayed as a % relative luminescent units (RLU) to PSV A at 0 mpi. Results represent three independent biological repeats with error bars indicating standard deviation. One-way ANOVA with Bonferroni correction for multiple comparisons was used for statistical analysis ($p < 0.0001$= ***, $p < 0.001$=** and $p < 0.01$=*). GraphPad Prism 5.0 software (CA, USA) was used. All Env clones were functional as PSV entry efficiency was higher than the background control when no PSV was added to the cells. Infection with PSV generated by transfecting HIV-1 pNL4.3luc+E-R- backbone with an empty vector was used as negative control for infectivity and subtracted from all RLU values.

Of the fourteen clones tested, PSV entry efficiency of seven differed in at least one cell line (Table 3.3). PSVs 37c3 and 137c3 relative entry efficiency was different in Affinofiles than the other two cell lines, suggesting that the low levels of CCR5 might be influencing entry of these clones. When Affinofiles were induced to express high levels of CD4 and varying levels of CCR5 (2,000 molecules/cell to ~31,000 molecules/cell molecules/cell), the ranking of PSVs from high to low entry efficiency was consistent across all CCR5 levels, suggesting that CCR5 levels were not influencing relative PSV entry efficiency (Figure 3.5).

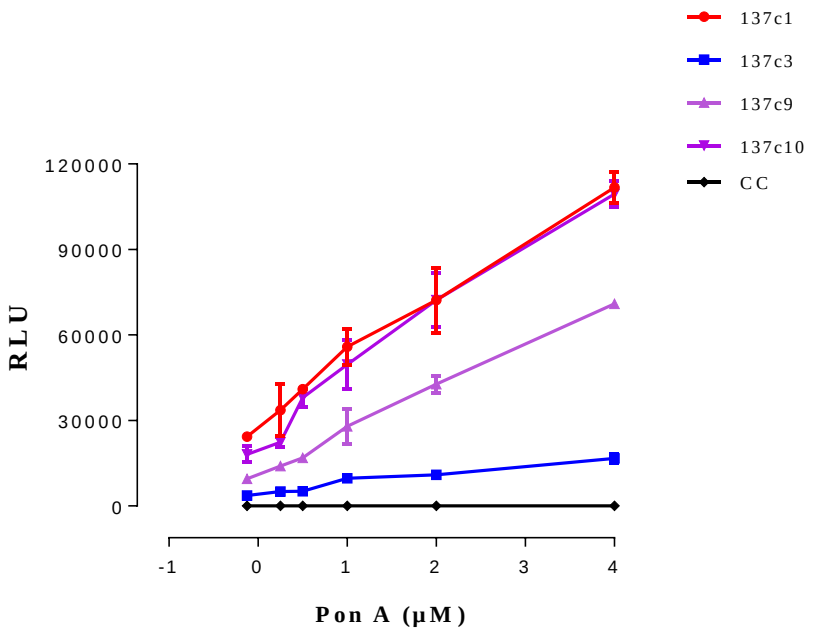


Figure 3.5. Pseudovirus entry efficiency relative to changes in CCR5 levels. Affinofile cells were induced to express high CD4 levels (4 ng/µl Minocycline) and varying levels of CCR5 using concentrations of Ponesterone that ranged from (0.13 to 4 µM). Pseudovirus (PSV) of CAP137 clones (c1, c2, c9 and c10 represented by circles, square, tringle, and upside-down triangles, respectively) was normalised to the equivalent of 100 ng p24 and entry was measured by relative light units (RLU). One biological repeat is indicated with error bars representing standard deviation of triplicate values.

If PSV entry efficiency differed in only TZM-bl cells then HIV backbones might be influencing Env entry efficiency. Only one clone, 137c10 had higher entry efficiency than other clones when tested in TZM-bl cells, suggesting that this clone might be influenced by

differences between pSG3delenv and pNL4-3.Luc.R^{-}E^{-}. The relative entry efficiency of the rest of the clones varied seemingly randomly across cell lines, suggesting that other physiological factors might play a role. It has been shown that differences in Env N-glycosylation has a discordant impact on virus transmission and infectivity depending on cell type (Shen et al., 2014). Correlation analysis showed a significant association (p = 0.049) between PSV entry efficiency in T-lymphocyte-like Affinofiles and TZM-bl cells (Figure 3.6), suggesting that the TZM-bl cell line can be used as a model for *in vivo* Env entry efficiency. However, differences between Affinofile (CD4^{high}/CCR5^{low}) cells and T lymphocytes might impact variation in entry efficiency of PSV clones.

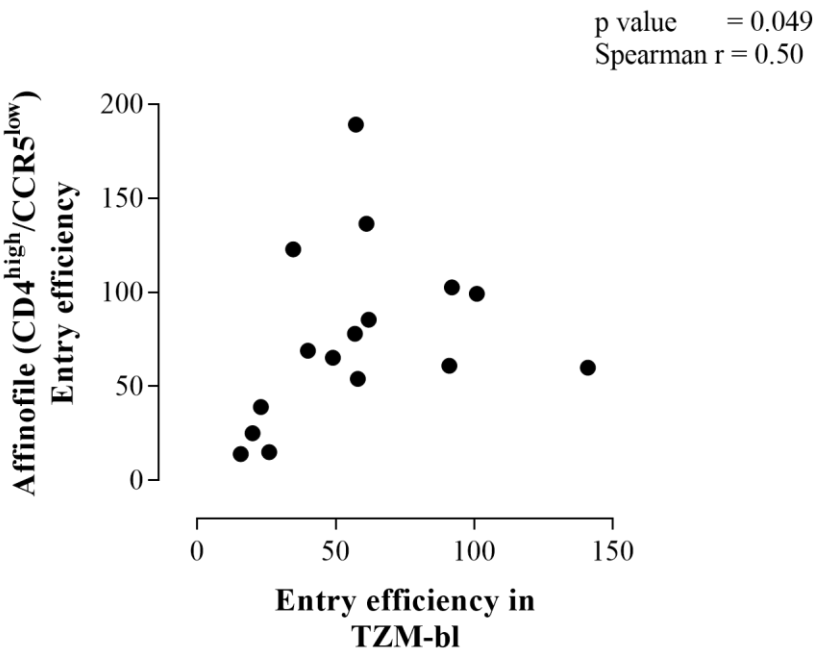


Figure 3.6. Correlation between PSV entry efficiency of TZM-bl and T-lymphocyte like Affinofile cells. Correlation analysis between the relative PSV Env entry efficiency in TZM-bl and T-lymphocyte was determined. To compare the entry efficiency of PSVs from independent experiments, RLU values were normalised to the background control when PSV was not added. Correlation co-efficient (r) was determined as indicative of a negligible, weak, moderate, high and very high relationship based on previous reports (Hinkle et al., 2003). p value of < 0.05 was considered significant. GraphPad Prism 5.0 software (CA, USA) was used.

3.4.4 Pseudovirus Fusion Capacity

Previous studies showed that viruses became less susceptible to entry inhibitors over time and this was associated with a significant increase in viral fusogenicity and infectivity (Etemad et al., 2009; Repits, 2005; Sterjovski et al., 2007). The fusion inhibitor, T-20, binds to the HR1 domain of gp41 and blocks host-viral membrane fusion (Reeves et al., 2005). In our study, increased entry efficiency for CAP137 and CAP267 seemed to be associated with changes within gp41, suggesting that gp41 might be an important determinant of Env function. As T20 inhibits gp41-mediated viral-host membrane fusion, T20 IC50 can be a surrogate marker for PSV fusion capacity (Cavrois et al., 2014; Reeves et al., 2005). The better the Env clone can mediate fusion the greater the concentration of T-20 needed to inhibit PSV entry and thus the higher the IC50, and vice versa. PSVs will be more susceptible to T-20 inhibition if Env has poor fusion capacity (Cavrois et al., 2014; Etemad et al., 2009; Lobritz et al., 2007; Reeves et al., 2002). Contrary to CAP137, TZM-bl entry efficiency of CAP37 and CAP84 PSVs did not increase over 12 months of infection but as Env fusogenicity might be more sensitive to changes in gp41 (Diaz-Aguilar et al., 2013) it might be easier to detect variation in fusion capacity than entry efficiency. We thus determined whether changes in gp41 altered T20 susceptibility without influencing Env entry efficiency.

CAP37

In CAP37, the fusion capacity between the Env clones varied 15 – 100 % but there was no overall change in the fusion capacity of variants infecting CAP37 over time. PSVs representing AB/B variants at 12 mpi (37c7) was ~7-fold more sensitive to T20 compared to virus A at enrolment (37c1) which could explain the low entry efficiency of this PSV. Overall, changes in PSV T20 IC50 reflected Env entry efficiency in TZM-bl cells, with both phenotypes fluctuating over 12 months infection (Figure 3.7A).

CAP84

In CAP84, virus A at 0 mpi (84c1) had a higher T-20 IC50 value compared to the recombinant virus AB at 12 mpi (84c4) (Figure 3.7B). This was similar to the difference in PSV entry efficiency.

CAP137

Two clones were generated at enrolment for CAP137 representing virus A: 137c1 and 137c2. The latter Env was 3.6-fold more susceptible to T-20 than 137c1, similar to the 2-fold difference in entry efficiency between the two clones. Sequence analysis revealed that they differed by 33 amino acids (a.a) in gp120 and 4 a.a in gp41. Moreover, 26/33 a.a were within V1/V2, previously reported to impact Env fusogenicity and thus Env entry (Cavrois et al., 2014). Therefore, these polymorphisms might contribute to the stronger fusion capacity of 137c1 PSV (Figure 3.7C).

The recombinant variants AB at 12 mpi (137c9 and 137c10) had higher fusion capacity compared to virus A at 0 mpi by 4- and 3-fold, respectively, mimicking entry efficiency data. However, even though PSV 137c10 entered TZM-bl cells (Table 3.2) more efficiently than 137c9, there was no significant difference in T-20 IC50 between 137c9 and 137c10 and, in fact, 137c9 was more fusogenic than 137c10. Although, sequence analysis showed that 137c9 and 137c10 varied mostly in gp41 (9 a.a in the HR2 region, 8 a.a in MPER and one in LLP-1 region) these polymorphisms might not only affect Env viral-host cell membrane fusion but other structural or phenotypic properties such as the interaction between gp41 and gp120 during entry (Etemad et al., 2009). It has previously been suggested that the influence of gp41 on entry efficiency and fusogenicity is highly clone specific (Yi et al., 2015). However, overall, variation in fusion capacity over the course of infection tracked that of PSV entry efficiency, suggesting that gp41 could be a major determinant in the entry efficiency of CAP137 PSV.

CAP267

Sequence analysis and mutagenesis suggested that five a.a. changes within gp41 were responsible for the 4-fold higher entry efficiency of 267c6 at 12 mpi compared to 267c2 at enrolment (Figure 2.12, Chapter 2). The fusion capacity of 267c6 PSV was also higher than that of 267c2, (28-fold higher), suggesting that the five a.a. that mapped to the cytoplasmic tail (CT) region (two a.a between Kennedy region and LLP-2, one in LLP-3 and two in LLP-1), might be influencing the fusion capacity of Env and potentially entry efficiency and *in vivo* frequency. Comparison between gp41 of CAP267 sequences aligned to HXB2 indicated

that 2/5 a.a. were conserved with consensus carrying an Asparagine (N) at position 750 and an Arginine (R) at position 845. However, 267c2 had an Aspartic acid (D) at position 750 and a Serine (S) at position 845 that might have reduced Env fitness. When D750 and R845 reverted to consensus in 267c6, this might have rescued the fitness cost. Reversion has been shown to come with a gain of fitness, potentially contributing to higher fusogenicity of 267c6 (Crawford et al., 2007; Song et al., 2014). As far as we know, none of these amino acids have been documented to have a direct effect on Env entry efficiency but the CT and LLP1 regions do play a role in fusogenicity (Da Silva et al., 2013; Kalia et al., 2003). Differences in Env entry efficiency of some variants infecting CAP267 was not directly related to changes in gp41 function, similar to previous studies, as virus A at 12 mpi (267c5) had 1.5-fold higher fusion capacity than 267c6, but 2-fold lower PSV Env entry efficiency (Diaz-Aguilar et al., 2013). However, this difference in fusion capacity was not statistically significant.

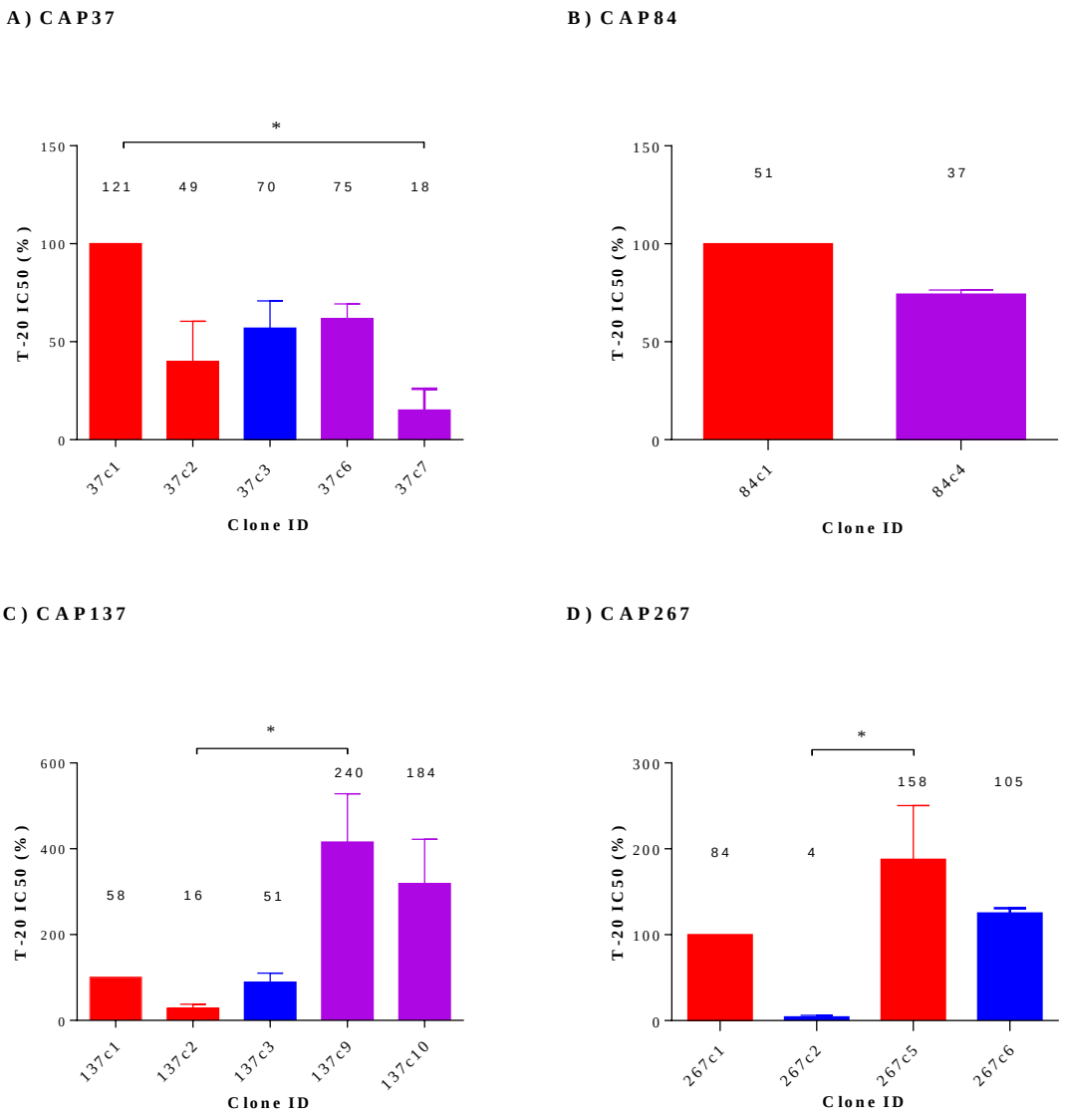


Figure 3.7. Pseudovirion sensitivity to the entry inhibitor, T20. The sensitivity to entry inhibitor (T-20) of pseudovirus (PSV) representing virus A, B and AB infecting A) CAP37, B) CAP84, C) CAP137, and D) CAP267 was compared over time (0, 2, and 12 mpi). The fusion capacity of PSV A, B and AB are indicated in red, blue and purple, respectively. The IC50 was compared between PSV and displayed as a relative luminescent units (RLU) to PSV A at 0 mpi (%). Results represent two independent biological repeats with error bars indicating standard deviation. The mean of T-20 IC50 of each Env is indicated at the top of each bar and used for statistical analysis. Significant difference is indicated by stars. One-way ANOVA with Bonferroni correction for multiple comparisons was used for statistical analysis ($p < 0.0001$= ***, $p < 0.001$=** and $p < 0.01$=*).

3.4.5 Association between PSV entry efficiency and fusion capacity

In some cases, T20 IC50 values did not support variation in entry efficiency between CAP137 and CAP267 clones similar to a previous study that showed the association between fusion and changes in entry efficiency was Env specific (Diaz-Aguilar et al., 2013). The authors suggested that there are mechanistic differences between Env entry efficiency and fusogenicity. Despite inconsistencies between PSV fusion capacity and entry efficiency in TZM-bl for some Env clones, analysis showed a significant correlation between fusion capacity and entry efficiency ($p = 0.02$, $r = 0.59$) (Figure 3.8). The moderate positive correlation suggests that any change in the membrane fusion step of viral entry might alter Env entry efficiency.

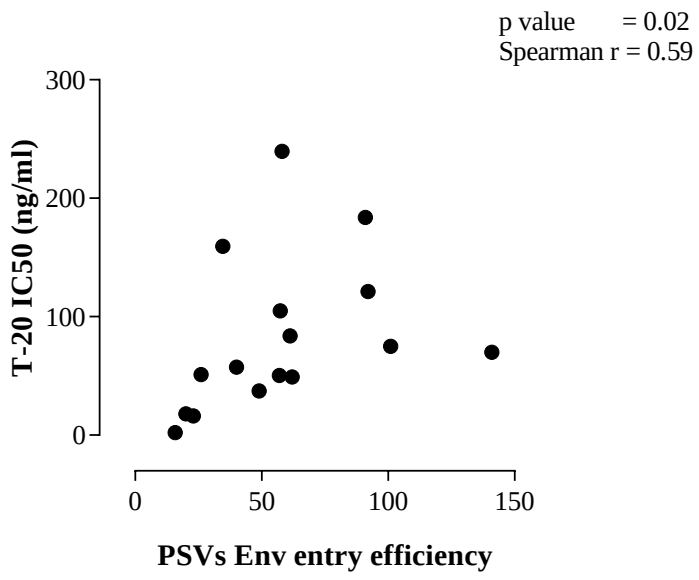


Figure 3.8. Correlation between pseudovirus entry efficiency and fusion capacity. The association between fusion capacity and corresponding Env entry efficiency of PSVs were analysed using Spearman r correlation test. Correlation co-efficient (r) was determined as indicative of a negligible, weak, moderate, high and very high relationship based on previous reports (Hinkle et al., 2003), p value of < 0.05 was considered significant. GraphPad Prism 5.0 software (CA, USA) was used for statistical analysis.

3.4.6 Envelope Expression and cleavage

HIV-1 infectivity is influenced by many factors, including the incorporation of functional trimers into virus particles, Env processing and cleavage of gp160 into gp120 and gp41 (Bachrach et al., 2005; Blay et al., 2007; Provine et al., 2009; Yuste et al., 2004). The outgrowth of variants infecting CAP137 and CAP267 was associated with increased PSV entry efficiency (Figure 2.9, Chapter 2) and fusogenicity potentially due to changes within gp41. Gp41 has been linked to Env cell-surface expression (Lebigot et al., 2001) which might lead to increased Env incorporation and enhanced entry efficiency (Bachrach et al., 2005). To determine whether the impact of gp41 on PSV entry efficiency was by enhancing Env expression levels, HEK293T cells were transfected, and the relative level of Env expression was compared by Western blotting. Env gp160 and gp120 band intensities were normalized to β-actin and virus A at enrolment was used as reference clone.

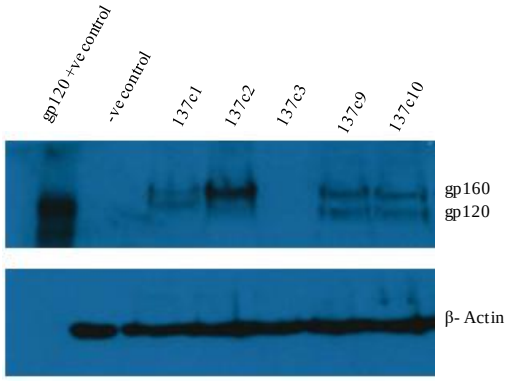
CAP137

Western blotting of 4/5 Envs showed that both cleaved and uncleaved gp160 were present in cell lysates (Asmal et al., 2011; Yu et al., 1993). Although bands representing 137c2 gp160 was 1.7-fold more intense than that of 137c1 gp160, gp120 levels were lower (Figure 3.9). There was no consistent relationship between gp160 expression and gp120 levels, supporting an earlier suggestion that Env over-expression may lead to inefficient cleavage of Env (Blay et al., 2007). As PSV entry efficiency is most likely influenced by gp120 density, we compared Env cleavage between clones.

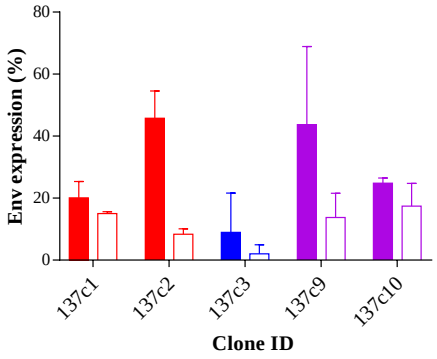
Virus B at 0 mpi (137c3) had the lowest cleavage of gp160 compared to all other Envs, suggesting that low gp120 incorporation into PSV could be driving its poor entry efficiency and low fusion capacity. On the contrary, Env expression/cleavage was not a contributing factor to the higher entry efficiency of PSV 137c10 compared to 137c9 in the TZM-bl cell line, as they had similar expression levels and gp160 cleavage. However, relative changes in cell-associated gp120 tracked PSV entry in Affinofile cells suggesting that, for some clones, cleavage of gp160 might be a determinant of Env entry efficiency. Despite a general trend between variation in PSV entry efficiency and Env cleavage, this relationship was not supported when comparing individual Envs using the TZM-bl reporter cell line.

CAP137

A)



B)



C)

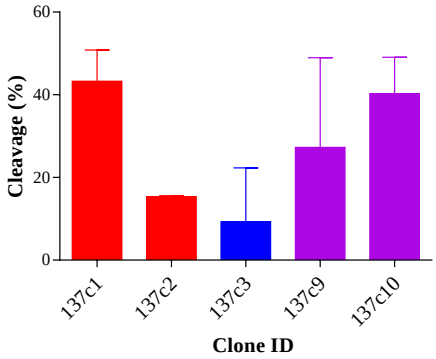


Figure 3.9. CAP137 Env Expression. Env expression of variants infecting CAP137 was determined by transfection of HEK293 T cells followed by quantitative Western blotting of cell lysates. A) One representative Western blot of two biological repeats. A total of 80 ng of protein was loaded per well and anti-gp120 antibody (AIDS Reagent Programme, ARP) and a conjugated anti-sheep antibody (Santa-Cruz Biotechnology) were used for detection. B-actin was used as a loading control and gp120+ is a recombinant purified HIV-1 gp120 (HIV-1 IIIB gp120 Recombinant Protein) from (ARP) was used as a positive control. Untransfected cells were used as a negative control (-ve control). B) Densitometry analysis using gel analyzer programme. Cell-associated gp160 and gp120 expression levels were normalized to β-actin. Results represent two independent biological repeats with error bars indicating standard deviation. Viruses A, B and AB are indicated with red, blue and purple bars, respectively. Gp160 and gp120 are indicated by solid and empty bars, respectively. C) Cleavage level of cell-associated gp120 relative to total Env (gp120/(gp160+gp120) *100) with viruses A, B and AB indicated by red, blue and purple bars, respectively. Kruskal-Wallis test with Dunn's multiple comparisons test was performed to compare cell-associated gp120 processing/cleavage level.

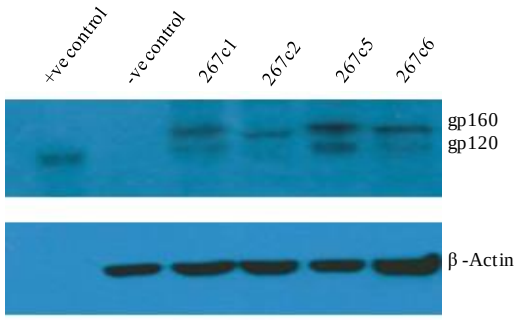
CAP267

Both 267c2 gp160 and gp120 bands migrated slightly lower than other Env clones which could be due to variation in Env glycosylation. Sequence analysis showed that 267c2 carried 26 PNGs, 6 less than 267c1 and 2 less than 267c6 which might have influenced its migration. There seemed to be a consistent relationship between gp160 and gp120 expression levels for all Envs (Figure 3.10). Env 267c5 had the highest expression: 1.3-fold and 1.9-fold better than 267c1 gp160 and gp120, respectively. This variation coincided with changes in PSV fusion capacity but not entry efficiency as 267c1 entered TZM-bl cells more efficiently. There were 18 a.a differences between 267c1 and 267c5 of which 4 mapped to the CT of gp41 previously reported to impact Env expression and incorporation (Jiang and Aiken, 2007). Virus B, 267c2, had the lowest gp120 expression, with levels 3-fold lower than 267c1, suggesting that expression could play a role in the low PSV entry efficiency and fusogenicity of this Env. Over time, virus B gp120 expression increased until 267c6 had 1.7-fold higher expression than 267c2, suggesting that point mutations within Env had increased expression and cleavage over the course of infection (Figure 3.10). Comparison between gp160 cleavage and Env entry efficiency suggested that the more cell-associated gp120, the higher the PSV entry. However, this relationship did not hold true for 267c6 that had the highest PSV entry efficiency but not the highest cell-associated gp120.

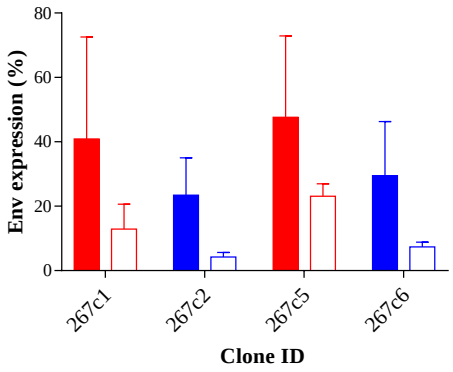
Correlation analysis showed a moderate positive correlation between the percentage of gp160 cleaved to gp120 and PSV entry efficiency ($r = 0.65$). However, this correlation was not statistically significant ($p = 0.07$) (Figure 3.11). Similarly, although the pattern of association between Env cleavage and fusogenicity was very similar, the moderate correlation was not significant ($p = 0.06$, $r = 0.65$). The lack of significance could be due to the samples size. Overall, this data suggests that the better gp160 is cleaved, the better fusion and thus entry (Bachrach et al., 2005; Blay et al., 2007; Provine et al., 2009; Yuste et al., 2004). Therefore, the variation observed in entry efficiency of most Env could be due to expression levels and cleavage, possibly by influencing the number of functional trimers incorporated into PSV.

CAP267

A)



B)



C)

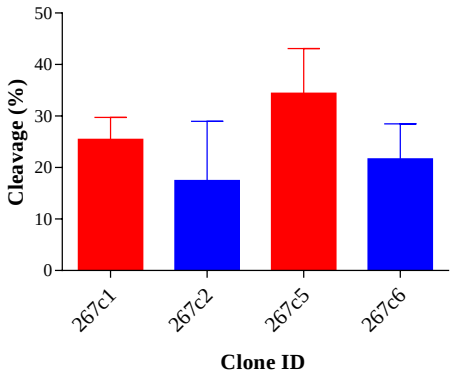


Figure 3.10. CAP267 Env Expression. Env expression of variants infecting CAP267 was determined by transfection of HEK293 T cells followed by quantitative Western blotting of cell lysates. A) One representative Western blot of two biological repeats. A total of 80 ng of protein was loaded per well and anti-gp120 antibody (AIDS Reagent Programme, ARP) and a conjugated anti-sheep antibody (Santa-Cruz Biotechnology) were used for detection. β-actin was used as a loading control and gp120+ is a recombinant purified HIV-1 gp120 (HIV-1 IIIB gp120 Recombinant Protein) from ARP was used as a positive control. Untransfected cells were used as a negative control (-ve control). B) Densitometry analysis using gel analyzer 2010 programme. Cell-associated gp160 and gp120 expression levels were normalized to β-actin. Results represent two independent biological repeats with error bars indicating standard deviation. Virus A and virus B are indicated with red and blue bars, respectively. Gp160 and gp120 are indicated by filled and empty bars, respectively. C) Cleavage level of cell-associated gp120 relative to total Env (gp120/(gp160+gp120) *100) with viruses A and B indicated by red and blue bars, respectively. Kruskal-Wallis test with Dunn's multiple comparisons test was performed to compare cell-associated gp120 processing/cleavage level.

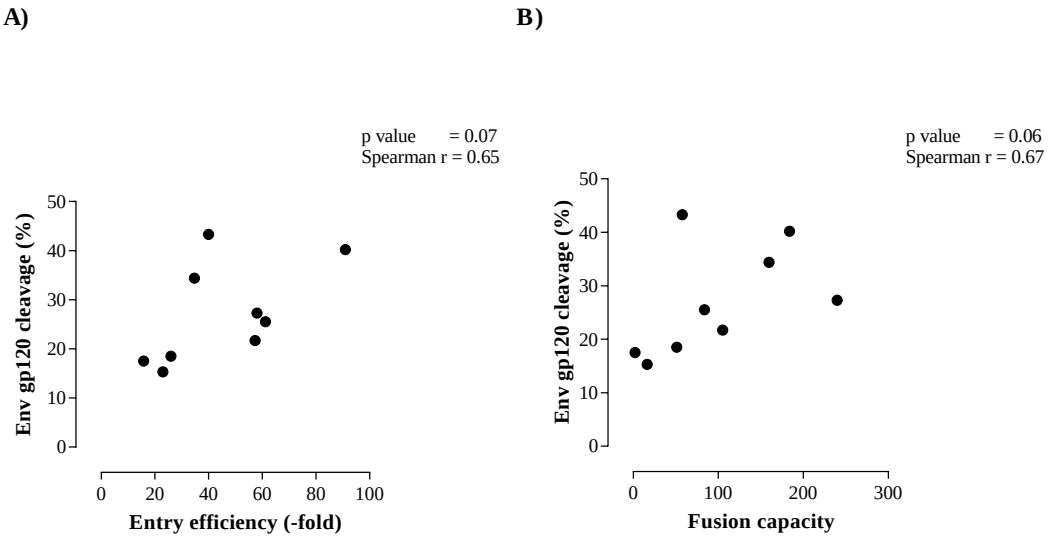


Figure 3.11. Correlation of gp160 cleavage with Env entry efficiency and fusion capacity. Correlation analysis between percentage cleavage of CAP137 and CAP267 gp160 into gp120 in the cell line and A) the relative PSV entry efficiency in TZM-bl and B) fusion capacity was performed using Spearman r correlation test. To compare the entry efficiency of PSVs from independent experiments, RLU values were normalised to the background control when PSV was not added. PSV fusion capacity was measured as the IC50 of T-20 inhibitor in TZM-bl. Level of cell-associated gp120 relative to total Env (gp120/(gp160+gp120) *100) was determined. Correlation co-efficient (r) was determined as indicative of a negligible, weak, moderate, high and very high relationship based on previous reports (Hinkle et al., 2003), p value of < 0.05 was considered significant. GraphPad Prism 5.0 software (CA, USA) was used for statistical analysis.

3.4.7 Env incorporation

Bachrach et *al*. (2005) suggested that slight increases in Env incorporation into viral particles profoundly increased infectivity (Bachrach et al., 2005). However, incorporated Env consisted of immature, uncleaved gp160, gp120 monomers, gp120-gp41 trimers and gp41 (Moore et al., 2006) with only incorporation of mature gp120 and gp41 trimers required for viral infectivity (Lodermeyer et al., 2013). It was suggested that density of mature trimers is dependent on Env expression and varies widely across HIV variants (Chertova et al., 2002).

We thus compared the relationship between Env expression, incorporation into viral particles and PSV entry efficiency.

CAP137

Western blot analysis could not resolve incorporated 137c1, 137c2 (in both biological repeats) and 137c9 (in one biological repeat) Env into clear bands corresponding to gp160 and gp120, most likely due to overloading and N-glycosylation (Figure 3.12). As virus infectivity is determined by gp120 incorporation (Bachrach et al., 2005; Blay et al., 2007; Stieh et al., 2015), we were unable to assess the relationship between gp120 incorporation and PSV entry efficiency. Env 137c3 was resolved into gp160 and gp120 bands and had the lowest levels of PSV-associated gp120 which could explain its poor ability to enter TZM-bl cells. However, PSV 137c10 with the highest entry efficiency of all the clones had similar PSV-associated gp120 levels to137c3, suggesting that for these clones, gp120 density was not influencing PSV entry.

CAP267

Despite the variation in Env gp160 and gp120 incorporation, there was no difference in the total amount of Env incorporated between virus A and virus B over time (data not shown). PSV 267c1, had the highest level of gp120 incorporated compared to the other clones (Figure 3.13). The ratio of gp160:gp120 decreased over time with 267c1 PSV carrying 4-fold greater gp120 than gp160 while an equal number of gp160 and gp120 seemed to be incorporated into 267c6 PSVs. Furthermore, 267c1 PSV-associated gp120 was 2.5-fold greater than 267c6 PSV. As 267c6 PSV entered TZM-bl cells with the highest efficiency compared to the other clones, it is unlikely that Env incorporation influenced PSV entry efficiency in TZM-bl.

As incorporation of gp120 into PSVs are essential for infectivity we tested whether cell-associated gp120 levels correlated with gp120 density on the surface of PSVs. Correlation analysis did not show any significance (Figure 3.14). Next, correlation analysis was performed between gp120 incorporation, and corresponding Env entry efficiency in TZM-bl using Spearman r correlation test. No significant correlation was found between gp120 incorporation and Env entry efficiency ($p = 0.3$, $r = 0.6$), suggesting that gp120 incorporation

did not impact entry efficiency (Figure 3.15). However, because we were unable to determine gp120 for two clones of CAP137, further analysis with a larger sample size is required.

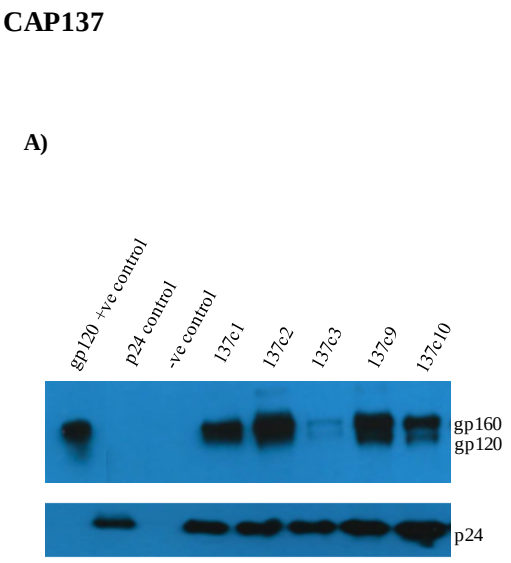


Figure 3.12. CAP137 Env incorporation into PSVs. Env incorporation of variants infecting CAP137 was determined by transfection of HEK293 T cells followed by quantitative western blotting of cell lysates. A) One representative Western blot of two biological repeats. Gp120 was detected using Anti-gp120 antibody (AIDS Reagent Programme, ARP) and a conjugated anti-sheep antibody (Santa-Cruz Biotechnology) were used for detection. P24 was detected using anti-p24 antibody (ARP432). p24+ and gp120+ are recombinant purified HIV-1 proteins (HIV-1 IIIB p24 Recombinant Protein and HIV-1 IIIB gp120 Recombinant Protein) obtained from ARP and used as a positive control. Viral particles without *env* were used as a negative control (-ve control). B) Densitometry analysis using gel analyzer 2010 programme. PSV gp160 and gp120 incorporation levels were normalized to p24 of corresponding virus. Results represent two independent biological repeats with error bars indicating standard deviation. Virus A, B and AB are indicated with red, blue, and purple colour, respectively. Gp160 and gp120 are indicated by filled and empty bars for each virus, respectively. Kruskal-Wallis test with Dunn's multiple comparisons test was performed to compare gp120 incorporation only.

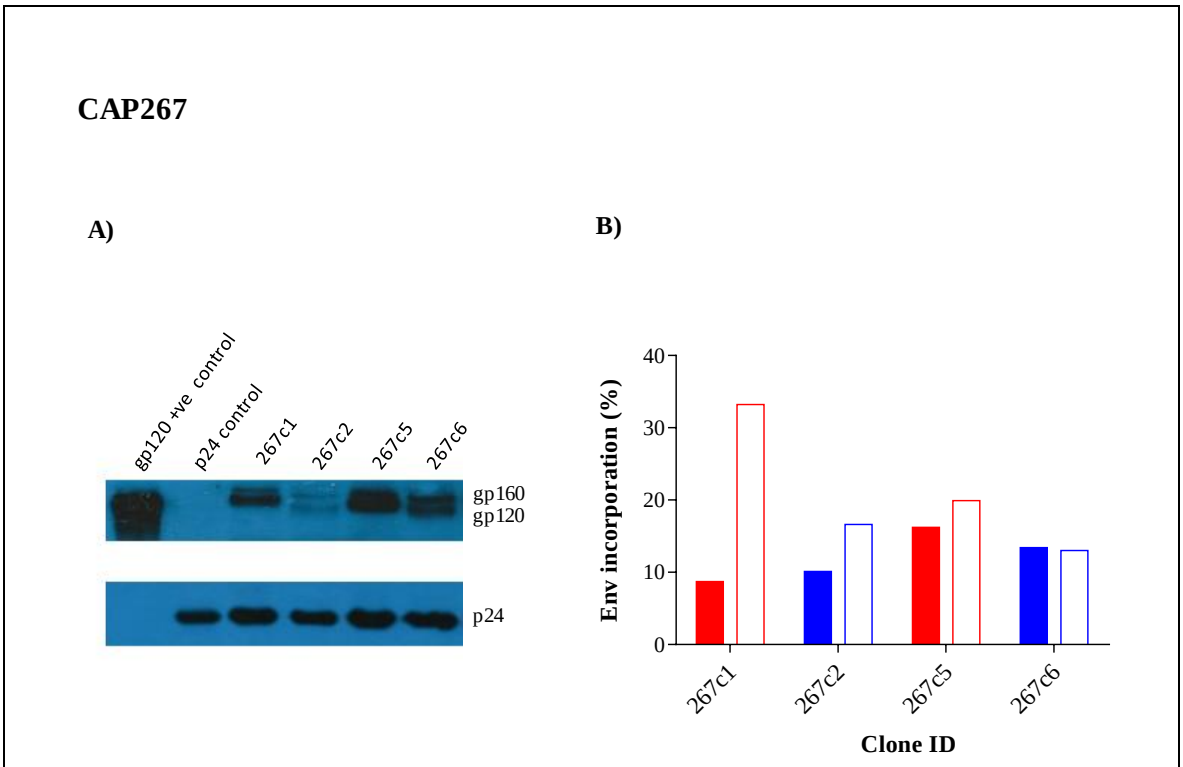


Figure 3.13. CAP267 Env incorporation into PSVs. Env incorporation of variants infecting CAP267 was determined by transfection of HEK293 T cells followed by quantitative Western blotting of cell lysates. A) One representative Western blot of two biological repeats. Gp120 was detected using anti-gp120 antibody (AIDS Reagent Programme, ARP) and a conjugated anti-sheep antibody (Santa-Cruz Biotechnology). P24 was detected using anti-p24 antibody (ARP432). p24 control and gp120 +ve are recombinant purified HIV-1 proteins (HIV-1 IIIB p24 Recombinant Protein and HIV-1 IIIB gp120 Recombinant Protein) obtained from ARP and used as a positive control. B) Densitometry analysis using gel analyzer 2010 programme. PSV gp160 and gp120 incorporation levels were normalized to p24 expression of corresponding PSV. Result represent one biological repeat. Virus A and B are indicated with red and blue bars, respectively. Gp160 and gp120 are indicated by filled and empty bars for each virus, respectively.

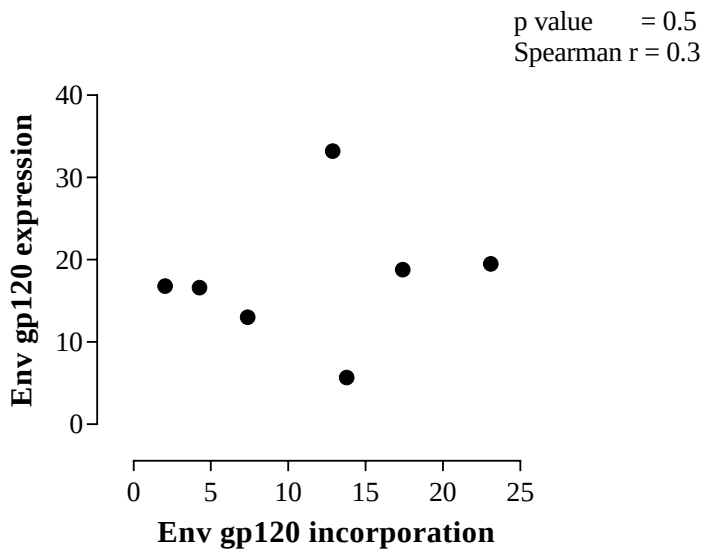


Figure 3.14. Correlation between Env gp120 expression and gp120 incorporation.
Correlation analysis between levels of cell associated gp120 and gp120 incorporated into pseuodvirions of CAP137 and CAP267 was performed using Spearman r correlation test. To compare the entry efficiency of PSVs from independent experiments, RLU values were normalised to the background control when PSV was not added. Gp120 incorporation band intensity was normalized to p24 and the mean of two independent biological repeats was determined. Correlation co-efficient (r) was determined as indicative of a negligible, weak, moderate, high and very high relationship based on previous reports (Hinkle et al., 2003), p value of < 0.05 was considered significant. GraphPad Prism 5.0 software (CA, USA) was used for statistical analysis.

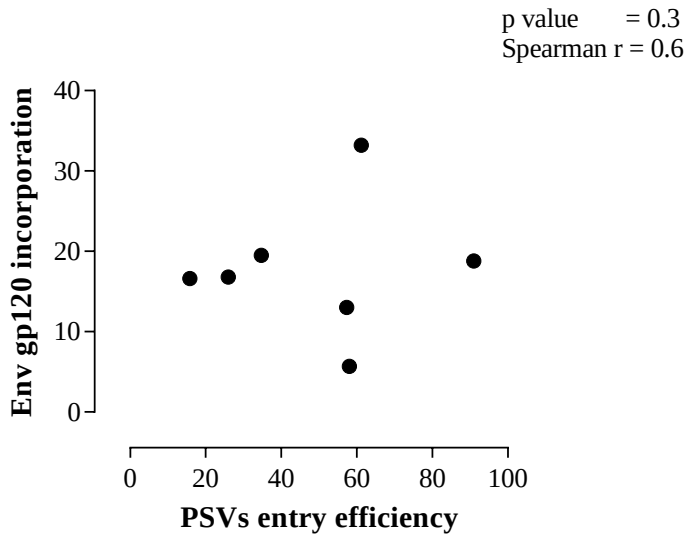


Figure 3.15. Correlation between Env entry efficiency and gp120 incorporation.
Correlation analysis between PSV entry efficiency of TZM-bl and Env gp120 incorporation level of CAP137 and CAP267 was performed using Spearman r correlation test. To compare the entry efficiency of PSVs from independent experiments, RLU values were normalised to the background control when PSV was not added. Gp120 incorporation band intensity were normalized to p24 and the mean of two independent biological repeats was determined. Because gp120 could not be resolved, three clones of CAP137 were excluded from the analysis. Correlation co-efficient (r) was determined as indicative of a negligible, weak, moderate, high and very high relationship based on previous reports (Hinkle et al., 2003), p value of < 0.05 was considered significant. GraphPad Prism 5.0 software (CA, USA) was used for statistical analysis.

3.4.8 Characterisation of Env fitness determinants

Analysis of Env phenotypic change over 12 months of infection indicated that, in general, Env fusogenicity was associated with variation in overall PSV entry efficiency. This relationship supported the apparent selection of changes in gp41 over the course of infection for viruses infecting CAP137 and CAP267. To confirm the importance of gp41 function, we determined whether sequence changes were not only linked to changes in entry efficiency but also Env fusogenicity, expression and incorporation (Cavrois et al., 2014; Lebigot et al., 2001; Reeves et al., 2005).

3.4.8.1 Fusion capacity of Env determinants

CAP84 and CAP137 variants with recombination in gp41 seemed to be selected at 12 mpi although only CAP137 PSV entry efficiency increased over time. Gp120/gp41 chimeras of CAP84 PSVs had enhanced entry efficiency, however, also suggesting that gp41 was important for changes in PSV entry. Although there was no obvious selection of CAP267 gp41 recombinants, chimeras suggested that gp41 could play a role in PSV entry efficiency increasing over time. As we had shown that Env fusogenicity was an important determinant of PSV entry efficiency, we wanted to confirm whether the PSV entry efficiency of the gp120/gp41 chimeras was similarly influenced by Env fusogenicity.

CAP84

Although 84c1 and 84c4 had similar entry efficiency, PSV of 84c1/c4 chimera with gp41 of virus 84c4, had 4-fold higher entry efficiency than PSV 84c1, suggesting that gp41 could play a role in the entry efficiency of this chimera. However, PSVs of both chimeras 84c1/c4 and 84c4/c1 had similar fusion capacity to each other and only 0.5-fold lower than wild-type Env clones (Figure 3.16A). This difference was not statistically significant suggesting that fusion capacity was not the major factor accounting for differences in Env entry efficiency of CAP84 chimeras.

CAP267

When comparing virus B variants infecting CAP267, PSV 267c6 at 12 mpi had 4.2-fold higher entry efficiency than 267c2 at 0 mpi. When the gp41 region was swapped between the two Envs, the PSV, carrying the gp41 of 267c6 had 4.5-fold higher entry efficiency, suggesting that gp41 could be driving the changes in Env entry efficiency and thus potentially viral outgrowth. Chimera 267c2/c6 also had enhanced fusion capacity similar to 267c6 while PSV 267c6/c2 fusion capacity was similar to virus B at enrolment 267c2 (Figure 3.16B). Therefore, the five a.a. changes within the CT of gp41 might be responsible for the increase in fusion capacity as well as the entry efficiency of virus B over time. Therefore, point

mutations within the CT of gp41 could be the major Env determinant for enhancing Env entry efficiency of this viral population.

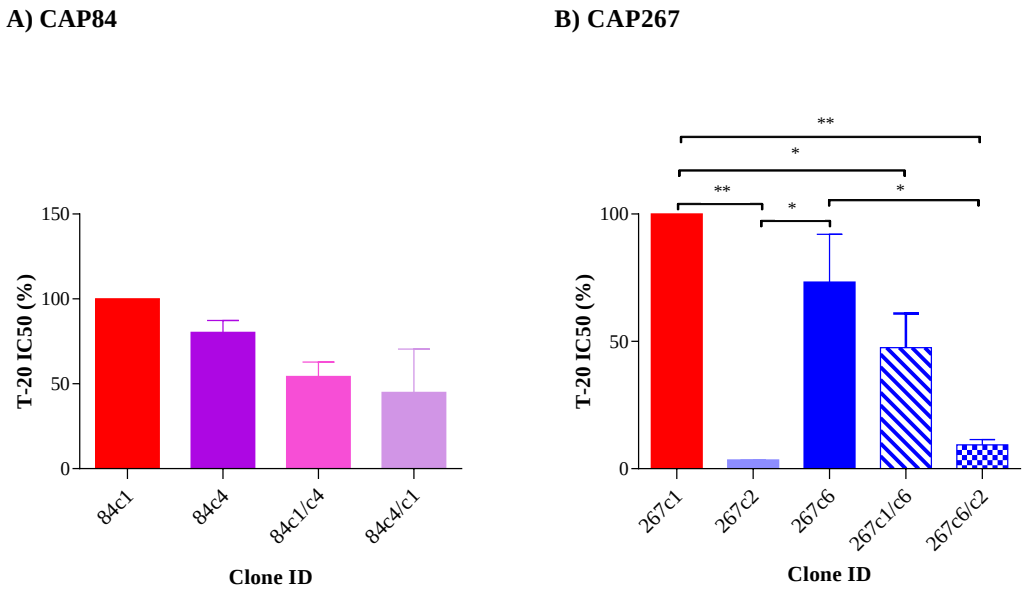


Figure 3.16. Fusion capacity of CAP84 and CAP267 chimeras. A) The fusion capacity of the two chimeras generated from 84c1 and 84c4. Cap84 virus A (84c1) at enrolment and AB (84c4) Env clone at 12 mpi are indicated in red and purple, respectively. The chimera 84c1/c4 comprises of the gp120 region of 84c1 and the gp41 of 84c4 and is indicated by a pink bar. The other chimera 84c4/c1 comprises of gp120 of 84c4 and gp41 of 84c1 and is indicated by a purple bar. Fusion capacity was compared to 84c1. B) The fusion capacity of the two chimeras generated from 267c2 and 267c6. Virus B at enrolment (267c2) is indicated by a light blue bar and virus B at 12 mpi (267c6) shown by a blue bar. The fusion capacity was compared to virus A at enrolment (267c1) indicated with a red bar. The first chimera (267c2/c6) comprises the gp120 of virus B at enrolment and gp41 of virus B at 12 mpi and indicated with a blue striped bar. The second one (267c6/c2) comprises of gp120 of virus B at 12 mpi and gp41 of virus B at enrolment indicated by blue checkered bar. One-way ANOVA with Bonferroni correction for multiple comparisons was performed using Prism 7.0. Data represent the average of two independent biological repeats with standard deviation ($p < 0.0001$= ***, $p < 0.001$=** and $p < 0.01$=*).

3.4.8.2 Determine the potential mechanism of CAP267 Env fitness determinant

Fusogenicity and PSV entry efficiency could be driven by levels of Env expression and incorporation into viral particles. Therefore, quantitative Western blotting of cell and pseudoviral lysates were performed to determine whether the five a.a. in gp41 was influencing fusion capacity by modulating Env expression and incorporation into viral particles. Visually, Western blotting of cell lysates of 267c6/c2 indicated very poor expression of gp120 and densitometry analysis could not detect the faint band corresponding to the molecular weight of gp120 (Figure 3.17A). However, Western blotting indicated that PSV-associated 267c6/c2 gp120 levels were only slightly lower than other PSVs (Figure 3.18).

Gp160 and gp120 expression of the chimera 267c2/c6, comprising gp41 of virus B at 12 mpi (267c6), was 1.6-fold and 1.1-fold, respectively, higher than that of 267c6 and had 2-fold higher gp160 and gp120 expression compared to 267c2 (Figure 3.17B). However, the percentage of gp160 cleaved into gp120 and the fraction of PSV-associated gp120 (Figure 3.18) did not vary between this chimera and 267c2 and 267c6 (Figure 3.17C). Overall, it would seem that the gp41 of 267c6 was not influencing Env entry efficiency by changing the density of PSV- associated gp120.

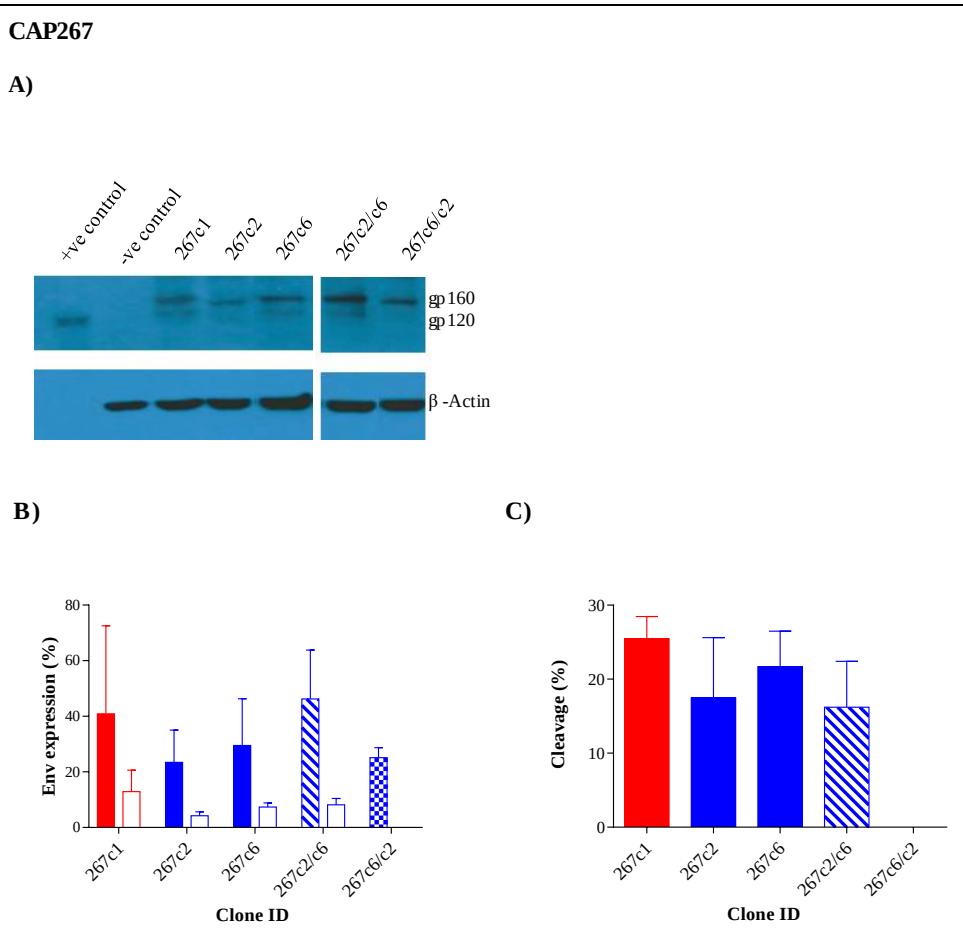


Figure 3.17. Env Expression and cleavage of CAP267 chimeras. Env expression of CAP267 chimeras was determined by transfection of HEK293 T cells followed by quantitative Western blotting of cell lysates. A) Representative Western blot of two biological repeats. A total of 80 ng of protein was loaded per well and anti-gp120 antibody (AIDS Reagent Programme, ARP) and a conjugated anti-sheep antibody (Santa-Cruz Biotechnology) were used for detection. β-actin was used as a loading control and gp120+ is a recombinant purified HIV-1 gp120 (HIV-1 IIIB gp120 Recombinant Protein) from (ARP) was used as a positive control. Untransfected cells were used as a negative control (-ve control). B) Densitometry analysis using gel analyzer 2010 programme. Cell-associated of gp160 and gp120 expression levels were normalized to β-actin. Results represent two independent biological repeats with error bars indicating standard deviation. Virus A at 0 mpi and virus B at 12 mpi are indicated with red and blue bars, respectively. The first chimera (267c2/c6) comprises the gp120 of virus B at 0 mpi and gp41 of virus B at 12 mpi and is indicated with a blue striped bar. The second chimera (267c6/c2) comprises of gp120 of virus B at 12 mpi and gp41 of virus B at 0 mpi indicated by blue checkered bar. Gp160 and gp120 are indicated by filled and empty bars, respectively. C) Cleavage. Level of cell-associated gp120 relative to total Env (gp120/(gp160+gp120) *100). Kruskal-Wallis test with Dunn's multiple comparisons test was performed to compare gp120 processing/cleavage level.

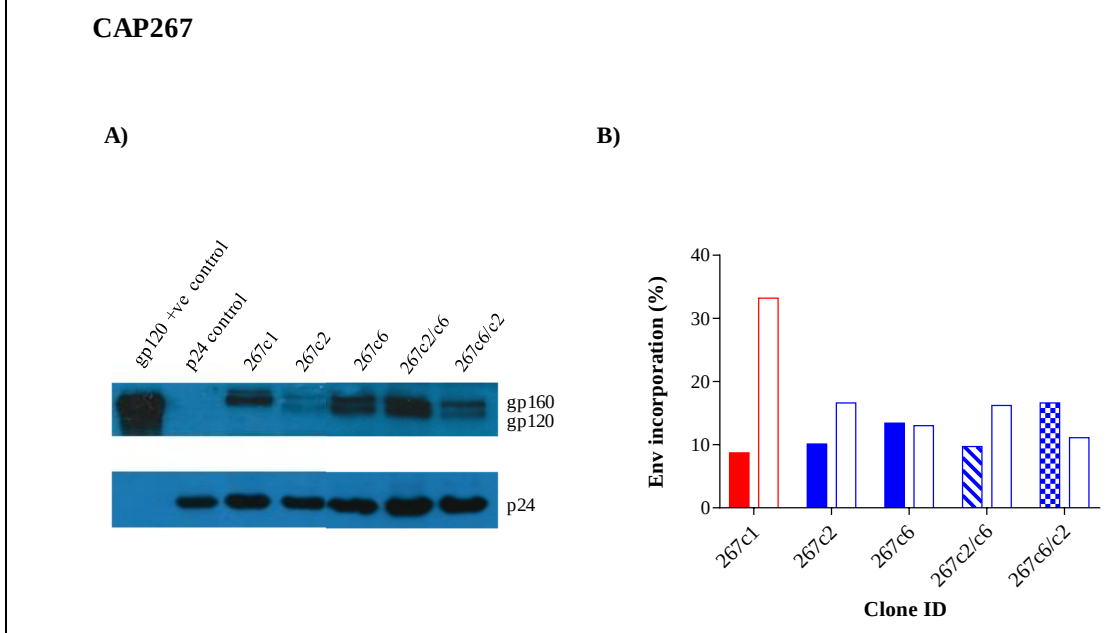


Figure 3.18. CAP267 Env incorporation of the chimeras. Env incorporation of CAP267 chimeras was determined by transfection of HEK293 T cells followed by quantitative Western blotting of pseudovirus lysates. A) Representative Western blot of two biological repeats. Anti-gp120 antibody (AIDS Reagent Programme, ARP) and a conjugated anti-sheep antibody (Santa-Cruz Biotechnology) were used for detection. P24 was detected using anti-p24 antibody (ARP432). p24+ and gp120+ are recombinant purified HIV-1 proteins (HIV-1 IIIB p24 Recombinant Protein and HIV-1 IIIB gp120 Recombinant Protein) obtained from ARP and used as positive controls. B) Densitometry analysis using gel analyzer 2010 programme. PSV gp160 and gp120 incorporation levels were normalized to p24 expression of corresponding virus. Results represent one biological repeats. Virus A and B are indicated with red and blue bars, respectively. The first chimera (267c2/c6) comprises the gp120 of virus B at 0 mpi and gp41 of virus B at 12 mpi is indicated with a blue striped bar and the second chimera (267c6/c2) comprises of gp120 of virus B at 12 mpi and gp41 of virus B at 0 mpi is indicated by a blue checkered bar. Gp160 and gp120 are indicated by filled and empty bars for each virus, respectively.

3.4.8.3 Determine the potential mechanism of CAP137 Env fitness determinant

A potential N-glycan at position 332 in gp120 has been linked to escape from BCN Ab (Moore et al., 2012). Analysis of CAP137 Env indicated that a PNG at position 332 appeared in 137c10 at 12 mpi. As 137c10, a representative of virus AB/B had the highest entry efficiency we determined the impact of a PNG at this site on Env function. Deletion of the PNG at N332 reduced PSV 137c10-sdm entry efficiency and to determine how a potential N-glycan at this position could influence Env function, we compared gp120 expression and incorporation into PSVs. On average, there was no difference in both overall Env expression (gp160+gp120) and gp120 levels of 137c10-sdm compared to wild type (Figure 3.19). Furthermore, although the incorporation of gp120 into 137c10-sdm PSVs tended to be lower than wild-type (1.5-fold) this relationship did not reach significance (Figure 3.20). Therefore, it is unlikely that the increase in entry efficiency observed for 137c10-sdm was due to a PNG at 332 enhancing Env expression and/or gp120 incorporation.

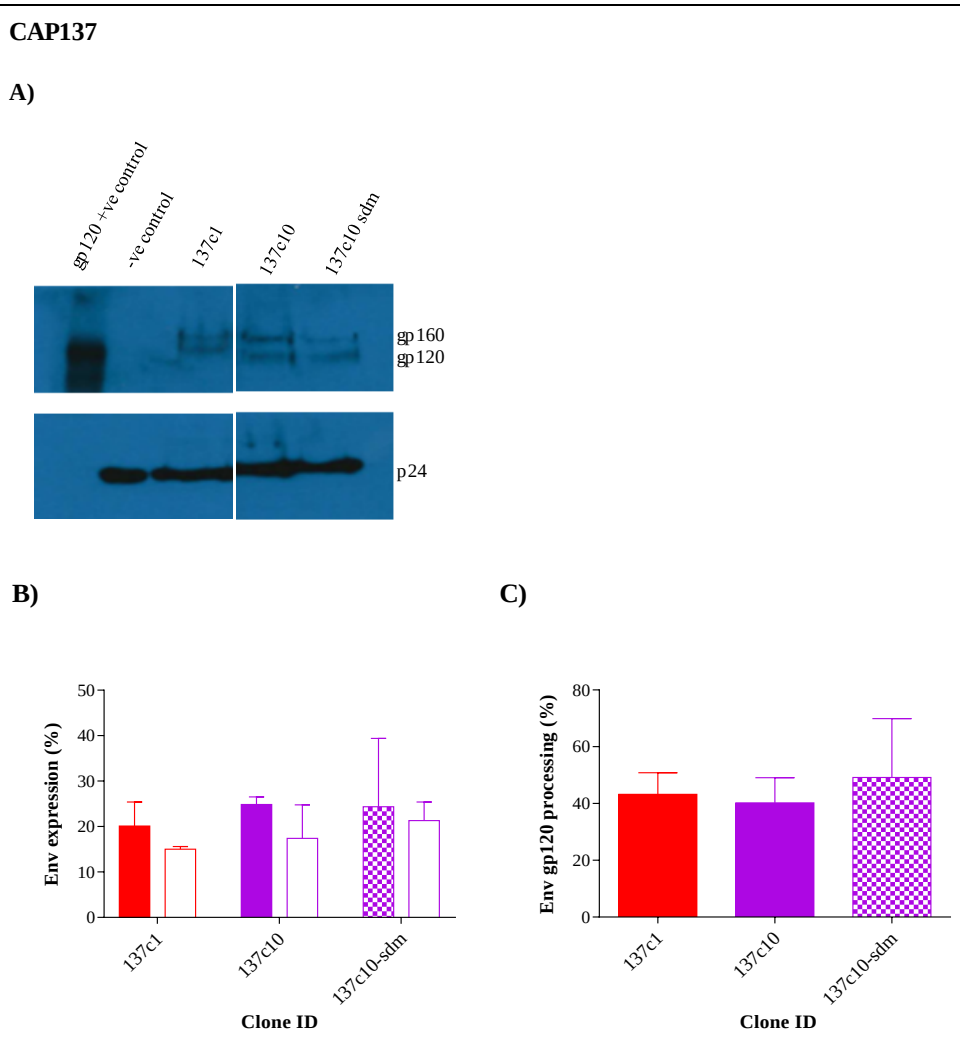


Figure 3.19. CAP137 Env Expression. Env expression of CAP137 wild type and CAP137-SDM with the N332S mutation was determined by transfection of HEK293 T cells followed by quantitative Western blotting of cell lysates. A) Representative Western blot of two biological repeats. Anti-gp120 antibody (AIDS Reagent Programme, ARP) and a conjugated anti-sheep antibody (Santa-Cruz Biotechnology) were used for detection. B-actin was used as a loading control and gp120+ is a recombinant purified HIV-1 gp120 (HIV-1 IIIB gp120) from ARP was used as a positive control. Untransfected cells were used as a negative control (-ve control). B) Densitometry analysis of Env expression. Cell-associated gp160 and gp120 levels were normalized to β-actin. Results represent two independent biological repeats with error bars indicating standard deviation. Gp160 and gp120 are indicated by solid and empty bars, respectively. C) Cleavage of gp160 was measured by determining cell-associated gp120 relative to total Env (gp120/(gp160+gp120) *100). Virus A (137c1), AB (137c10) and AB mutant (137c10-sdm) are indicated by red, purple and checkered purple bars, respectively. Kruskal-Wallis test with Dunn's multiple comparisons test was performed to compare gp120 processing level.

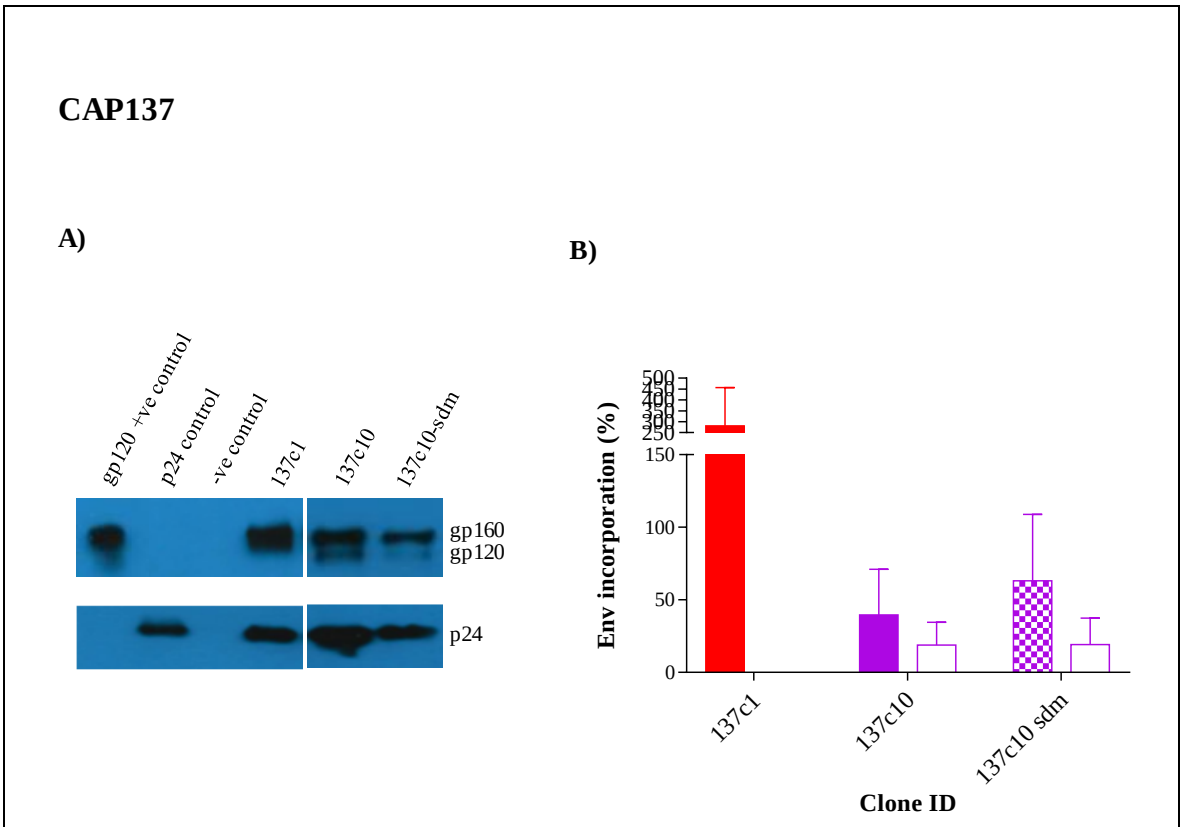


Figure 3.20. CAP137 mutant incorporation. Env incorporation of CAP137 mutant clone (137c10-sdm: N332S) and its wild type Env clone was determined by transfection of HEK293 T cells and ultra-centrifugation followed by quantitative Western blotting of pelleted pseudovirus lysates. A) One representative Western blot of two biological repeats. A total of 80 ng of protein was loaded per well and anti-gp120 antibody (AIDS Reagent Programme, ARP) and a conjugated anti-sheep antibody (Santa-Cruz Biotechnology) were used for detection. P24 was detected using anti-p24 antibody (ARP432). p24+ and gp120+ are recombinant purified HIV-1 proteins (HIV-1 IIIB p24 and HIV-1 IIIB gp120) obtained from ARP and used as a positive control. Viral particles without Env were used as a negative control (-ve control). B) Densitometry analysis of pseudovirus-associated Env. PSV gp160 and gp120 incorporation levels were normalized to p24 of the corresponding virus. Results represent two independent biological repeats with error bars indicating standard deviation. Virus A (137c1), AB (137c10) and AB mutant (137c10-sdm) are indicated by red, purple and checkered purple bars, respectively. Gp160 and gp120 are indicated by solid and empty bars for each virus, respectively. Unpaired nonparametric Student's t-tests, Mann Whitney test was performed to compare gp120 incorporation level between the wild type and mutant.

3.4.9 Summary of results

Table 3.4. Overall relationship between PSV entry efficiency, in vivo viral outgrowth and Envelope function and processing

Participant ID	Association of Increased Entry efficiency with variant outgrowth		Association of Increased Entry efficiency over 12 mpi with:				
	Over 12 mpi	Within time points	Fusion capacity	Expression	Gp160 cleavage	Gp120 incorporation	Fitness determinant
CAP37	No	No	Yes	ND	ND	ND	N/A
CAP84	No	Yes	Yes	ND	ND	ND	gp41/N339
CAP137	Yes	*Yes	*Yes	Yes*	Yes*	No	gp41/ N332
CAP267	Yes	Yes	*Yes	No*	No*	No	gp41

* Association was not apparent for all clones

Table 3.5. Phenotypic characterisation of Envelope fitness determinants

			Phenotype of Fitness determinant			
Participant ID	Fitness determinants	Mutants generated	Entry efficiency	Fusion Capacity	Expression	Incorporation
CAP84	Gp41	Chimeras: 84c1/c4 and 84c4/c1	Yes	No	ND	ND
	PNG at N339	introduction of N339 to 84c1/c4	Yes	ND	ND	ND
		Deletion of N339 in 84c4	No	ND	ND	ND
Cap137	PNG at N332	Deletion of N332 in 137c10	Yes	ND	No	No
CAP267	5 a.a in gp41 CT and one a.a in C5	Chimeras: 267c2/c6 and 267c6/c2	Yes	Yes	No	No

3.5 Discussion

The study of *in vivo* competition between variants infecting dual infected individuals provides insight into the selection of variants with specific characteristics/phenotypes or biological properties of Env that provide advantages to the survival and pathogenesis of HIV-1. Understanding advantageous Env characteristics can provide targets for viral control and attenuation if not clearance. Therefore, this study describes the evolutionary changes in Env phenotypes, such as receptor/coreceptor preferences, fusion capacity, incorporation and expression, over the course of infection that might influence entry efficiency of Env clones, viral fitness and thus disease progression of dual infected individuals.

In this study PSV entry of the reporter cell line TZM-bl was used as a measure of Env fitness and U87-CD4+CCR5+ and Affinofile cells were used for phenotypic characterisation. However, lab-adapted cell lines might not be the ideal replacement for T-lymphocytes, the natural host of HIV-1. Thus, the first objective of the study was to compare PSV entry efficiency of all three cell lines. We hypothesised that if PSVs entered TZM-bl cells, U87-CD4+CCR5+ cells and Affinofiles induced to express T-lymphocyte-like levels of CD4 and CCR5 with the same relative efficiency, they would most likely infect T-lymphocytes (*in vivo*) similarly.

CD4 density was similar for TZM-bl and U87-CD4+CCR5+ cells and approximately 3-fold higher than Affinofile $CD4^{high}/CCR5^{low}$ cells. There was greater variation in CCR5 levels with Affinofile $CD4^{high}/CCR5^{low}$, U87-CD4+CCR5+ and TZM-bl expressing 2000, 6000 and 11000 molecules/cell, respectively. Although most PSVs entered the three cell types similarly, there were instances when clones infected one cell type better than another. Comparison between cell types indicated that these differences could be due to CCR5 levels, HIV backbone differences as well as other unidentified cell-specific characteristics. However, there was a significant correlation between entry of Affinofile $CD4^{high}/CCR5^{low}$ and TZM-bl cells, suggesting that overall, TZM-bl cell entry was a good indicator of *in vivo* Env fitness. Furthermore, seeing that CCR5 levels, and Env fitness, differed between individuals (Garg et al., 2016) utilising PBMCs to compare Env entry might also have disadvantages.

The efficiency by which HIV-1 utilises CD4 and CCR5 during entry impacts viral infectivity, pathogenesis and thus disease progression. Studies have reported that some R5 viruses evolve to macrophage tropism and can infect cells with low CD4 levels, thus expanding the number of cells permissive to HIV infection (Blaak et al., 2000; Chikere et al., 2014; Joseph et al., 2014). Our data showed that all PSVs, isolated at 0 mpi and at 12 mpi, required high levels of CD4 to enter cells, consistent with previous studies that showed no difference in receptor/coreceptor utilisation over time (Parrish et al., 2012; Ping et al., 2013). It is possible that evolution to macrophage tropism is restricted to specific compartment such as CNS where the neutralizing antibody pressure is low (Beauparlant et al., 2017; Ping et al., 2013).

Previously, it was hypothesised that the genetic bottleneck at transmission selected for single variants that were R5 and T-tropic as it provided the transmitted founder with a survival advantage (Ping et al., 2013). However, variants transmitted to dual infected individuals were also R5 and T-tropic, suggesting that phenotypic selection was not restricted to single infection events. Overall, neither viral tropism nor cellular tropism contributed to the differences observed in Env entry efficiency over time and thus these Env phenotypes are unlikely to be the driving factor for variant outgrowth in dual infected individuals.

Our data showed that PSV entry efficiency was significantly influenced by the fusion capacity of Env for all participants ($p = 0.02$). Env fusogenicity of CAP37 and CAP84 tracked changes in PSV entry efficiency and therefore did not increase over time. Env fusion capacity increased over time in two participants, CAP137 and CAP267 consistent with enhanced PSV entry efficiency (Table 3.4). Increased resistance to T20 of longitudinally isolated viruses during chronic stages of infection was previously reported, suggesting that increased Env fusogenicity provided a survival advantage over the course of infection (Chatziandreou et al., 2012; Etemad et al., 2009; Repits, 2005).

CAP37, CAP84, CAP137 and CAP267 Env sequence analysis suggested that gp41 might be under positive selection. Furthermore, gp120/gp41 chimeras of CAP84 and CAP267 indicated that gp41 might be an important fitness determinant of PSV entry efficiency. The main function of gp41 is mediating viral-host membrane fusion however studies have shown that it's important in a number of other processes. The CT is involved with the viral cycle at

numerous steps including: Env incorporation, expression and stability of the gp41/gp120 heterodimer and viral fitness (Da Silva et al., 2013). Mutations in LLP1 of the CT decreased the fusion capacity of Env leading to decreased viral infectivity (Kalia et al., 2003). While some studies proposed that truncation of LLP1 stimulated changes of the helical structure of LLP-2 and/or changes in the structure of the Env ectodomain, (Jiang and Aiken, 2007; Kalia et al., 2003; Wyma et al., 2004), other studies suggested that the interactions of LLP1with LLP2 and the cell membrane (Wyss et al., 2005) and/or with regions of the gp41 ectodomain, particularly the HR1-HR2 6-helix bundle was responsible for changes in viral infectivity (Abrahamyan et al., 2005). The CT was also documented to interact with Gag during virus assembly and alter the conformation of gp41 which in turn impacted Env incorporation into virions and influenced viral infectivity (Davis et al., 2006).

For CAP137 and CAP267, a.a. changes within gp41 sequences mapped mainly to HR and LLP1/LLP3, respectively. As the C5 region only differed in one a.a. we proposed that the 5 a.a. in the CT of CAP267 Env influenced PSV entry efficiency over time (Chapter 2) and therefore CAP267 Env chimeras were tested for changes in fusion capacity. When gp41 was swapped between virus B at enrolment and virus B at 12 mpi Env fusogenicity increased, suggesting that Env fusion capacity was influencing PSV entry efficiency. Closer analysis of the CT indicated that two of the five a.a., N750 and R845, were highly conserved indicating positive selection pressure. However, 267c2, carried D750 and S845, suggesting that these a.a. might have lowered PSV entry efficiency and when they reverted to consensus in 267c6, entry efficiency increased (Song et al., 2014). Therefore, we showed that the CT of gp41 could be important for enhancing Env entry efficiency of viral population B over time in CAP267, consistent with previous studies (Jiang and Aiken, 2007; Kalia et al., 2003).

The entry efficiency of CAP37 and CAP84 decreased over time, similar to changes in Env fusogenicity, suggesting an association between the two phenotypes. However, when gp41 was swapped between CAP84 variants with different entry efficiency and fusion capacity the two chimeras had similar susceptibility to fusion inhibitor, suggesting that gp41 was not important for changes in Env fusogenicity for these clones. As the chimeras differed in Env entry efficiency but not fusogenicity, clone specific differences most likely influenced the effect of gp41 (Yi et al., 2015).

Previous studies reported that CT mutations influenced Env expression, processing, incorporation and virus infectivity (Bhakta et al., 2011; Kalia et al., 2003). In our study there was no association between PSV entry efficiency and fusogenicity with expression, and incorporation. Despite the moderate positive correlation between the cleavage of gp160 and both Env entry efficiency ($p = 0.07$, $r = 0.65$) and fusion capacity ($p = 0.06$, $r = 0.67$), the relationship was not statistically significant. Reports have been conflicting with some showing a correlation between gp120 incorporation and viral infectivity (Blay et al., 2007; Parrish et al., 2013) whereas another study did not (Jiang and Aiken, 2007). Bachrach *et al*., (2005) had suggested that Env density no longer influenced infectivity once levels increased above a certain threshold so it is possible that our PSVs carried excess Env (Bachrach et al., 2005). Viral populations consist of three subgroups: virus carrying 1) functional, 2) non-functional and 3) mixed functional and non-functional Env and the first group makes up the bulk of the population (Blay et al., 2007; Moore et al., 2006; Stieh et al., 2015). It is expected that poor Env cleavage would increase the level of incorporation of non-functional forms of Env (Blay et al., 2007; Herrera et al., 2005) which non-native Western blotting would not be able to distinguish. Therefore, even though 137c3 and 137c10 PSV carried similar gp120 levels, it is possible that the number of functional molecules incorporated per viral particle might differ between clones. However, in our study, there was no correlation between the percentage of cell-associated gp120 and the level of gp120 incorporated into PSVs. It is possible that changes in gp41 LLP1 altered incorporation of gp120 without influencing Env expression and/or processing as previously reported by Kalia *et al.*, (2003) (Kalia et al., 2003).

When we analysed the phenotypes of potential Env fitness determinants our data from the two chimeras of CAP267 supported the lack of association between gp120 incorporation and PSV entry efficiency as the density of gp120 did not vary between 267c2/c6, 267c2 and 267c6. Despite the PNG at 332 of CAP137 showing an effect on PSV entry efficiency, it had no significant impact on Env expression or Env incorporation but could play a role in other Env phenotypes, such as the affinity to CD4 or/and CCR5. Structural modelling (Figure 3.21 showed that this PNG (C3 region) was located near the V3 loop that plays a crucial role in coreceptor binding. Therefore, N332S could play a role in stabilizing the V3 loop and

modulating coreceptor binding and thus Env entry efficiency (Wang et al., 2013). However, the effect of this N-glycan varied between different HIV-1 variants as it caused significant disruption of Env function of virus YU2 but not of JR-FL (Lavine et al., 2012). Therefore, further mutational analysis of other clones should be carried out to confirm the influence of this PNG on viral fitness.

Furthermore, the PNG at 339 (C3 region) in CAP84 viruses that affected PSV entry efficiency for some clones, was also located in a region neighbouring the V3 loop involved in coreceptor binding. Therefore, introduction of a PNG at 339 could indirectly affect Env entry efficiency by causing a shift in the variable V3 loop, decreasing the binding affinity for CCR5 (Wang et al., 2013). However, further analysis is still required.

Conclusion

In summary, our data suggests that gp41 is an important determinant of Env function by modulating Env fusion capacity. Thus, Env fusion capacity plays a very important role in PSV entry efficiency and potentially *in vivo* viral outgrowth in dual infected individuals.

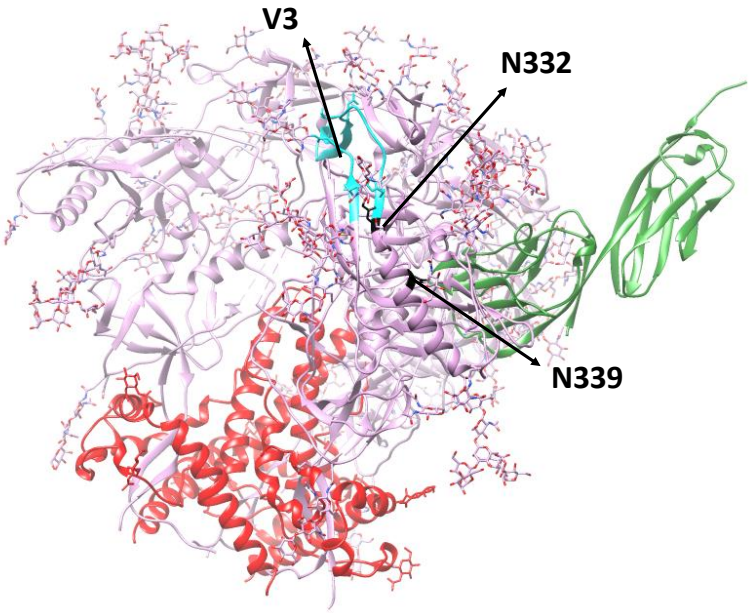


Figure 3.21. Structural modelling of potential N-glycans at positions N332 and N339. Homology model of a complex trimer of Env gp120 (in purple) and gp41 (in red) showing the positions of N332 and N339 relative to theV3 loop (in cyan colour), a CCR5 binding site, and the bridging sheet (in green). The structure was modelled using UCSF Chimera 1.11.2 software based on BG505 SOSIP.664 trimer in complex with the broadly neutralizing antibody 3BN117 (PDBe 5v8m) (Lee et al., 2017). EMDataBank: EMD-8644.

CHAPTER 4.

Understanding the Effect of Envelope on Viral Replication

4.1 Introduction

Viral load set-point and CD4 T cell decline are surrogates for disease progression and it has been suggested that viral genotype and thus HIV-1 traits, contribute to the heritability of viral loads (Bertels et al., 2017). Replication capacity (RC) of HIV-1 viruses was found to correlate with high viral load and disease progression (Selhorst et al., 2017; Williamson and Swanstrom, 2015). We showed earlier that in dual infected individuals, Env entry efficiency tend to play an important role in the outgrowth of recombinant viruses. Furthermore, there was a significant association between loss of CD4 count and PSV entry efficiency suggesting that Env was playing a role in disease progression. Although, PSV-based assays have been extensively used to estimate Env entry efficiency and suited for high throughput, it measures only a single cycle of viral replication (Etemad et al., 2009; Quiñones-Mateu and Arts, 2001).

Multiple cell culture assays have been used to assess viral replication capacity and fitness *in vitro* (Dykes and Demeter, 2007; Quiñones-Mateu and Arts, 2001). In multiple cycle assays, (i) the same culture can be co-infected with different variants in a competition assay or (ii) replication of viruses can be compared independently in parallel culture (Dykes and Demeter, 2007; Quiñones-Mateu and Arts, 2001). Growth competition assays were potentially limited by recombination between the two infecting/competing variants (Song et al., 2012) at high multiplicity of infection (MOI) (Ball et al., 2003; Lanxon-Cookson et al., 2013) and distinguishing between viruses usually required expensive methods of quantification such as deep sequencing or real-time PCR. On the other hand, parallel growth kinetics assays were less expensive and successfully used to evaluate Env function, phenotype and RC of HIV-1 variants in the absence (Chatziandreou et al., 2012; Etemad et al., 2009; Quiñones-Mateu and Arts, 2002; Selhorst et al., 2017; Weber et al., 2011) and presence of drugs (Selhorst et

al., 2017; Weber et al., 2011). The outcome of replication assays can be influenced by cell type and donor-derived peripheral blood mononuclear cells (PBMCs) is considered a more relevant measure of virus replication than cell lines that have adapted to laboratory conditions (Quiñones-Mateu and Arts, 2002; Salazar-Gonzalez et al., 2009; Troyer et al., 2005). Therefore, to confirm that Env influenced viral RC similar to that of PSV entry efficiency, chimeric infectious molecular clones (IMCs) were generated using HIV-1 NL4-3 proviral genome backbone and parallel replication in PBMCs monitored by p24 ELISA. Furthermore, we evaluated whether Env determinants influenced viral replication as these did not consistently affect Env entry efficiency, fusion, expression, cleavage and incorporation into PSVs.

4.2 Research Aim and Objectives

Aim

Determine the relationship between RC of variants isolated from dual infected individuals and *in vivo* outgrowth of variants over the course of infection and the effect of Env fitness determinants on viral RC.

Objectives

Objective 1: To determine the RC of chimeric IMCs and compare RC with *in vivo* outgrowth of variants and Env phenotypes.

Objective 2: Compare RC to PSV-entry efficiency to understand the role of Env in virus replication.

Objective 3: Describe the effect of Env determinants on RC.

4.3 Material and Methods

4.3.1 Samples

CAP137 and CAP267, both classified as rapid progressors were selected to generate IMCs from Env isolated at 0 mpi and 12 mpi. In addition, IMCs were also generated for the mutant clones 137c10sdm, 267c2/c6 and 267c6/c2.

4.3.2 Plasmids used in the assay

The shuttle plasmid pCMV-NL4-3-PBS→LTRΔGp160 with the HIV NL4-3 subtype B genome (without the 5'LTR) and the helper plasmids pCMV-NL4-3-LTR→Gag4 that carries the NL4-3 5′ LTR were generous gifts from Dr. Manish Sagar, Brigham and Women's Hospital, Harvard Medical School, US (Figure 4.1). Plasmid construction was as described before (Chatziandreou et al., 2012; Dudley et al., 2009). Briefly, the subtype B NL4-3 full-length genome was inserted into the multiple cloning site (MCS) of pRS315 using the yeast gap-repair homologous recombination (pRS315-NL4-3) prior to replacing the primer binding sequence (PBS) of NL4-3 5'LTR sequence with *URA3* gene to form pRS315-NL4-3Δ5'end to prevent the expression of the entire genome of NL4-3. Subsequently, the CMV promoter was amplified from pCDNA3 and inserted into the pRS315-NL4-35' to form pCMV-NL4-3-PBS-3'LTR, placing genes of NL4-3 under the control of the CMV promoter. Finally, *env* was replaced with the *URA3* gene amplified with primers that introduced a unique restriction enzyme *Pac*I site to form pCMV-NL4-3-PBS/LTRΔGp160. Yeast gap-repair homologous recombination was used to insert *env* to form the final shuttle vector. For constructing the helper plasmid, the yeast centromere sequence (CEN6) and LEU2 gene were amplified from pRS315 and inserted into MCS of pcDNA3 to generate pcDNA-Leu. The 5'LTR to gag sequence was also inserted into pcDNA-Leu to form CMV-NL4-3-LTR -Gag4 (Helper).

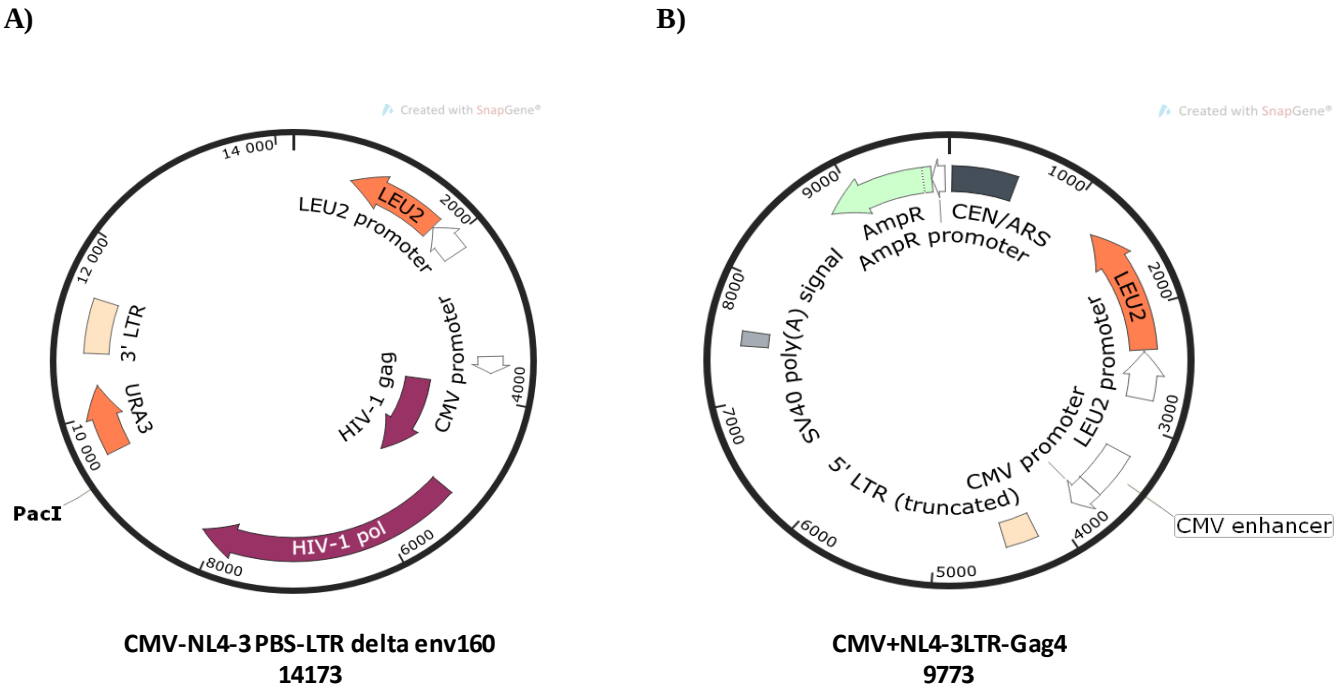


Figure 4.1. Map of shuttle vector and helper plasmid used in yeast recombination assay. A) Shuttle vector: pCMV-NL4-3-PBS/LTRΔgp160 in which the URA3 gene replaced the *env* gene in the NL4-3 genome with a unique restriction enzyme PacI allowing other *env* genes to be inserted into this plasmid using yeast homology recombination. B) Helper plasmid: pCMV-NL4-3-LTR→Gag4 comprising the sequence spanning the 5'LTR to *gag* under the control of the CMV promotor. Plasmid map was constructed using SnapGene viewer software.

4.3.3 Yeast recombination assay

In order to generate IMCs, yeast homologous recombination was used as described previously (Chatziandreou et al., 2012; Dudley et al., 2009).

4.3.3.1 Env gene amplification

Full-length HIV-1 *env* was amplified from 137c1, 137c2, 137c3, 137c9 and 137c10 (Table 2.1, Chapter-2). PCR was carried out using the Phusion Hot Start Polymerase kit (Thermo Scientific, USA) with the two primers, Env IF and Env IR (Appendix A) using the parameters in (Appendix A).

The PCR product (3 kB) was excised from the 1 % agarose gel and purified using the Wizard® SV Gel and PCR Clean-Up System following the manufacturer's instructions (Promega, USA). The purified PCR product was sequenced at Stellenbosch Central Analytical Facility using an ABI 3000 genetic analyser (Applied Biosystems, Foster City, CA, USA) and BigDye terminator reagents. The shuttle vector, pCMV-NL4-3-PBS/LTRΔGp160 (3 µg) was linearized with 3 µl of FastDigest *Pac*I restriction enzyme (1U/ µl), (Thermo Scientific, USA) and cleaned up with the Wizard® SV Gel and PCR Clean-Up Kit (Promega, USA) following the manufacturer's instructions.

4.3.1.2 Yeast gap-repair recombination

The yeast were transformed with the amplified *env* genes as previously described (Chatziandreou et al., 2012; Dudley et al., 2009). Briefly, yeast [(*Saccharomyces cerevisiae* Hanson (MYA-906), MAT α ade6 can1 his3 leu2 trpl URA3)] cells were made competent using 10 % glycerol and stored in 100 µl aliquots at -80 °C. The yeast cells were thawed at room temperature and centrifuged at 13000 rpm for 30 seconds and the supernatant was removed. A mixture of the linear *env* PCR product (1 µg/µl) and the linear vector (200 ng/µl) in a ratio of 5:1 (*env*:vector) that made up to a total volume of 74 µl with distilled water was added to the yeast cells. Transformation was performed by adding 240 µl of 50 % polyethylene glycol (PEG3350) (Sigma Aldrich, USA) (Appendix C), 10 µl of 10 mg/ml salmon sperm (Invitrogen), and 36 µl of 1M lithium acetate (Appendix C). The

transformation mixture was incubated at 30 °C for 30 min, followed by heat shock at 42 °C for 15 minutes. The cells were centrifuged for 30 seconds at 13000 rpm and resuspended gently with 100 µl of sterile distilled water before plating on Complete Supplement Mixture (CSM) – Leucine + 5 –Fluoroorotic acid (CSM-LEU + 5'FOA) selective plates (Appendix C). Only transformed yeast should grow as 5'FOA is toxic to cells expressing URA3. The plates were incubated at 30 °C for 3-4 days till colonies became visible. Yeast cells transformed with distilled water instead of DNA and/or yeast cells transformed with digested vector only were used as negative controls. Yeast cells transformed with PRS315 plasmid and plated on Leucine-only agar plates controlled for transformation efficiency.

4.3.3.2 Yeast plasmid extraction

To extract the recombinant plasmid, two to three colonies were selected from CMS- Leu/5-FOA selective plates and cultured in 2 ml of CMS-leucine broth (Appendix C) at 30 °C with shaking overnight. Zymoprep™ Yeast Plasmid Miniprep II kit from (Zymo Research, USA) was used to extract the recombinant plasmid from the yeast following the manufacturer's instructions. PCR amplification with two primers, *env*-y F and *env*-y R (Appendix A) was used to confirm the successful cloning of the *env* gene into the recombinant plasmid. These primers were designed specifically for the subtype C *env* used in this study. The PCR product (approximately 1 kb in size) was confirmed on 1% agarose gel. Moreover, the recombinant plasmids were sequenced at the Stellenbosch Central Analytical Facility using an ABI 3000 genetic analyser (Applied Biosystems, Foster City, CA, USA) and BigDye terminator.

4.3.3.3 Yeast plasmid transformation into bacterial competent cells

The commercial 10 Beta *E. coli* electro-competent cells from (New England BioLabs, UK) were transformed with a volume of 1-2 µl of yeast-extracted plasmids using a Gene Pulsar II electroporator (Bio-Rad, USA) following the manufacturer's instructions. After electroporation, the cells were incubated at 30 °C for 1 hour with SOC medium (provided by manufacturer) before plating on ampicillin selective Luria agar plates and incubated at 30 °C overnight. The PureYield™ Plasmid Miniprep System (Promega, USA) was used to extract the recombinant plasmid from the bacterial culture as per manufacturer's instructions.

4.3.4 Generation of Infectious Molecular Clones

HEK 293T cells were co-transfected with 2 µg of the recombinant plasmid (pCMV-NL4-3-PBS/LTR + *env*) and 2 µg of the helper plasmid (CMV NL4-3 Gag4) at a molar ratio of 1:1.8 using polyethyleneimine (PEI) (Sigma-Aldrich). Briefly, $4x10^5$ cells/ml were plated in 6 well plates and incubated overnight at 5 % CO_2, 37 °C before transfection. A ratio of 1:3 (DNA:PEI) was vortexed in 400 µl of DMEM (Sigma, Germany) for 10 seconds and incubated at ambient temperature for 15 min before adding dropwise to HEK 293T cells in 2 ml supplemented DMEM (Section 2.3.4.1, Chapter 2). The supernatant containing viral particles was collected after 48 hours, clarified through 0.45 um filter, and a volume of 600 µl was aliquoted into cryotubes and stored at -80 °C till used.

4.3.5 Virus Titre

An aliquot of viral stock was thawed at room temperature and serial diluted with supplemented DMEM medium in 1:3 or 1:5 dilution series in triplicate in 96 well plates in a total volume of 200 µl. A volume of 100 µl aliquot of each dilution was transferred to TZM-bl cells plated at a concentration of 10^4 cells per well in 100 µl supplemented DMEM media in 96 well plates (Costar; Promega). Cells were incubated at 37 °C, 5 % CO_2 for 48 hours. Volume of 150 µl of the medium was removed and 100 µl Bright-Glo Luciferase substrate-buffer mix (Promega, USA) was added to the TZM-bl cells, incubated for 2 min to lyse the cells before 100 µl were transferred into a white opaque 96 well plate and relative light units (RLU) measured by a GloMax-Multi Microplate Multimode Reader (Promega, USA). RLU readings of uninfected cells were considered as background signal and infection was considered successful if RLU was > 2.5-fold higher than background. In the case of titrating viruses expanded in PBMCs, Dextran (40 µl per 10 ml supplemented DMEM, 20 µg/ml final concentration) was added to the supplemented DMEM of TZM-bl prior to plating the cells. This was used to maximize virus infectivity. The 50 % tissue culture infectivity dose (TCID50) was calculated based on Reed-Muench method (Reed and Muench, 1938).

4.3.6 IMC Replication

4.3.6.1 Virus expansion

Ficoll-gradient centrifugation was used to isolate PBMCs from HIV-negative donors. Cells were maintained at 37 °C with 5 % CO_2 in RPMI 160 medium supplemented with 10 % Fetal Calf Serum (BiochromGmBH), 25 mM HEPES (Lonza, Basel, Switzerland), 1 U/mL penicillin and 1 µg/mL streptomycin (Whitehead Scientific) antibiotic mixture, and 2 mM L-glutamine (Lonza, Basel Switzerland). PBMC proliferation was stimulated with IL-2 (Gentaur, Belgium) (200 U/ml final concentration) and phytohemagglutinin-P (PHA-P) lectin (Remel, Thermo Scientific, MA, USA) (0.5 µg/ml final concentration) in RPMI 160 for 72 hours before virus expansion or virus replication.

In order to produce high virus titre, the recombinant viruses generated from HEK 293T cells were expanded in PBMCs. Activated PBMCs were plated in 6 well plates at a concentration of 5 x 10^6 cells per 2 ml RPMI supplemented with IL-2 only and 500 µl of each virus stock was added. Spinoculation was performed by centrifugation at 1500 g for 2 hours at room temperature to enhance virus infectivity. The infected cells were then transferred to T25-tissue culture flasks and medium was topped up to 5 ml supplemented RPMI 160 medium with IL-2 so that cells were at a final concentration of 10^6 cells per ml. The flasks were then incubated at 37 °C, in a 5 % CO_2 incubator for two weeks. Infected PBMCs were refreshed with 5 million activated PBMCs in fresh supplemented RPMI 160 medium with IL-2 every three days and virus supernatant was collected from day 7 and every three days thereafter. A volume of 1 ml of virus supernatant was aliquoted in cryovials and stored at -80 °C.

4.3.6.2 Replication Kinetic assay

PBMCs were isolated from HIV- negative donors and activated for 72 hours as previously described. PBMCs were plated at a concentration of 10^6 cells per ml in 96 well plates, and 300 TCID50 of each virus was used to infect the cells in triplicate in a final volume of 200 µl per well. Cells without virus were used as the negative control. Infection was enhanced by centrifugation at 1500 g for 2 hours. Infected cells were then incubated at 37 °C in a 5 % CO_2 incubator and 50 µl of virus supernatant was collected on day 7, 10, and 14 post-infection, replaced with fresh medium and stored at -80 °C for further analysis. Virus replication was monitored by measuring p24 using an in-house p24 ELISA technique (Alto-Biosystems) as previously described (Section 2.3.4.3, Chapter 2). Replication kinetics was

plotted as p24 ng/ml vs days post- infection, and two analytical methods were used to compare RC between IMC chimeras: (i) area under the curve (AUC): area under the p24 curve was estimated using GraphPad Prism version 5.0 (Chatziandreou et al., 2012) and (ii) slope of viral replication: slopes between RLU values at days 0 and 7, 0 and 10, 0 and 14 were calculated and averaged for each virus and the mean and SEM was determined of two independent measurements (Weber et al., 2011). NL4-3 HIV provirus was used as a positive control and cells without infection was used as a negative control.

4.3.6.3 PBMC donor selection

There was high variability between donors with some not able to sustain viral replication. Thus, to identify donors that were consistently permissive to HIV infection, PBMCs were isolated from four different HIV-negative donors and an aliquot was activated and immediately tested while the remaining were cryopreserved at -80 °C for future use. Activated cells were infected with 300 TCID50 of three viruses: two lab control viruses (NL4-3 and Bal) (Gifted by Prof. C. Williamson, UCT) and one of the chimeric IMCs generated for this study (137c10). Infected cells were incubated at 37 °C in 5 % CO_2 for 14 days and supernatants were collected as previously described. Viral replication was determined by p24 ELISA for day10 samples only.

4.3.7 Statistical analysis

IMC RC was compared using one-way ANOVA test with Bonferroni correction for multiple comparisons using GraphPad Prism version 5.0. p value of < 0.05 was considered as statistically significant.

4.4 Results

4.4.1 Generation of Infectious Molecular Clones

Chimeric IMCs were generated by cloning *env* into the HIV-1 NL4-3 proviral backbone to test the impact of Env on viral replication. The yeast gap repair technique was used to rapidly insert different *envs* into the pNL4-3 backbone as cloning by restriction enzyme digestion and ligation is limited due to the extreme heterogeneity between HIV variants (Dudley et al., 2009).

Previously, we showed that a recombinant variant infecting CAP137 and CAP267 virus B dominated at 12 mpi with concomitant increased PSV entry efficiency (Table 2.1; Figure 2.9, Chapter 2). The changes in variant frequency and corresponding PSV entry efficiency of these two participants suggested that Env, and more specifically, gp41 could play an important role in viral fitness. Therefore, we determined whether the influence of Env on PSV entry could be extrapolated to IMC RC.

Five IMCs were generated from CAP137, two representing virus A at 0 mpi, one representing virus B at 0 mpi, and two representing the AB viral population at 12 mpi (Section 2.3.3, Chapter 2). In addition, an IMC was also generated for the mutant clone 137c10sdm (Section 2.3.5.2, Chapter 2). For CAP267, four Env IMCs were selected, two representing A and B viral populations at 0 mpi and two representing the A and B viruses at 12 mpi. Additionally, IMCs were also generated for the two chimeric mutant constructs (267c2/c6 and 267c6/c2) discussed in section 2.3.5.1, Chapter 2.

4.4.2 Identification and isolation of responsive donor PBMCs

It was previously reported that *in vitro* HIV viral replication was influenced by variation between PBMCs from different donors making is difficult to determine whether differences between Envs were due to donor variation or intrinsic Env function (Spira and Ho, 1995). Furthermore, we found that some donor PBMCs did not support viral replication irrespective of whether CD8+ T cells were depleted or not (data not shown). Therefore, in order to select the most responsive donor PBMCs and obtain sufficient cells to test all Envs with the same

donor, we isolated PBMCs from four different HIV negative donors, generated aliquots for storage at -80 °C and tested the RC of two lab strain viruses (NL4-3 and Bal) and the replication competent chimeric IMC, 137c10.

The RC of NL4-3, Bal and 137c10 differed within and between donor PBMCs with cells isolated from donor-2 the least permissive (Figure 4.2). The RC of NL4-3 was consistently more robust than 137c10, even though they only differed in Env. The replication of NL4-3 and Bal was lower in donor-2 than that in the other three donors and 137c10 failed to replicate in these cells. Therefore, PBMCs isolated from donors 1, 3 and 4 were selected to compare IMC replication. However, in subsequent experiments we were unable to detect replication of our chimeric IMCS using donor-3 PBMCs and we will thus only report on two independent replication assays.

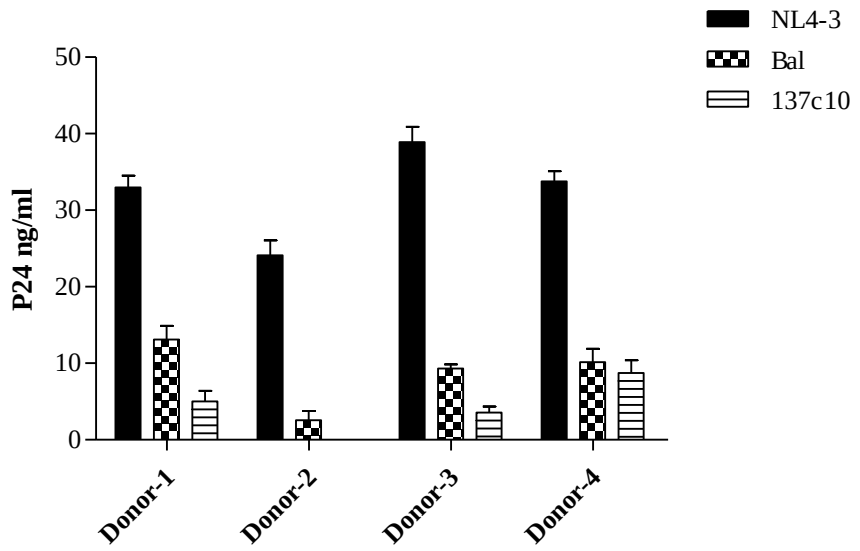


Figure 4.2. Replication in PBMCs from different donors. PBMCs from four donors were activated and infected with 300 TCID50 of two virus controls, NL4-3 and Bal and one experimental IMC, 137c10 in triplicate. Virus supernatant was collected and p24 concentration was measured at day10 using p24 ELISA. The mean p24 concentration is indicated with the error bars representing the standard deviation.

4.4.3 Replication of IMCs in PBMCs

In vitro parallel, non-competitive replication assays have been used to compare the fitness of Envs from heterosexual and homosexual transmission, drug users, longitudinal samples and variants from long-term non-progressors (LTNP) to those of progressors as well as between subtype C, subtype A and subtype B viruses (Bunnik et al., 2010; Chatziandreou et al., 2012; Claiborne et al., 2015; Etemad et al., 2009; Pernas et al., 2012; Selhorst et al., 2017). Here we generated chimeric IMCs carrying Env of longitudinal samples from dual infected individuals which represented viruses A, B and/or AB as they fluctuated over the course of infection of CAP137 and CAP267.

CAP137

Previous studies have used changes in p24 over time, the AUC (Chatziandreou et al., 2012), or the slope (Weber et al., 2011) to compare variation in replication fitness between variants. To determine whether these analytical approaches yielded the same results, we calculated changes in mean p24 concentration over time, the AUC of two independent experiments from two donors and the mean slope of the p24 curve.

IMCs representing the AB viral population at 12 mpi had the highest RC: 137c9 and 137c10 had 9-fold and 4-fold higher RC than virus A at 0 mpi, respectively (Figure 4.3A and B). This indicated that fitter viruses emerged over time (Figure 4.3). The RC of 137c9 was greater than the lab-adapted strain pNL4.3, suggesting high viral fitness. The same relationship between viruses was apparent when comparing the mean AUC with slight variation in the degree of statistical significance (data not shown).

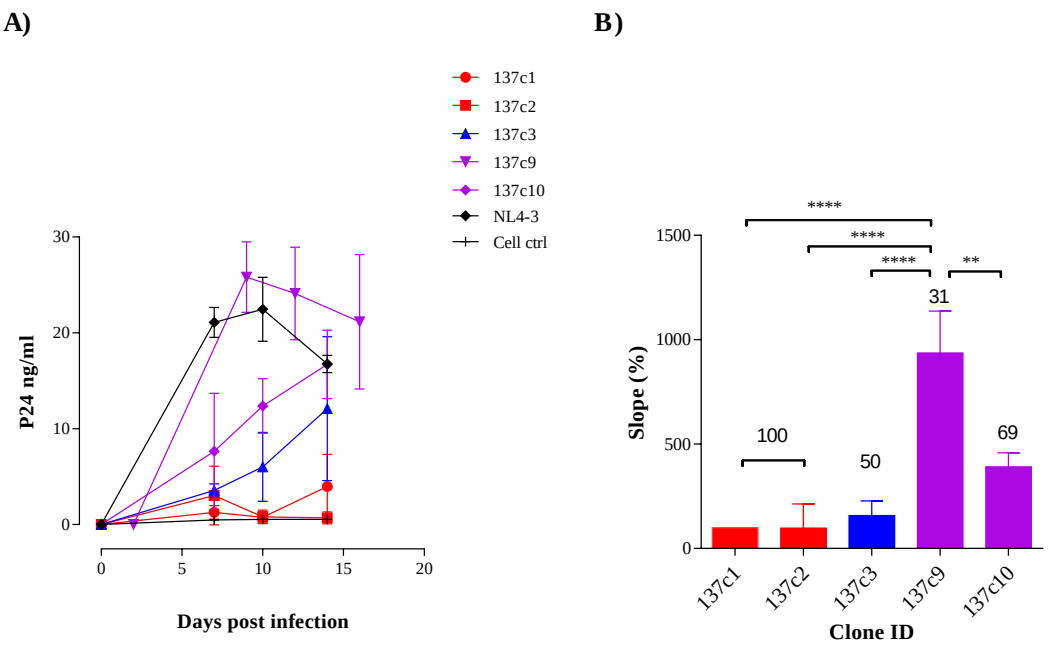


Figure 4.3. Replication of CAP137 IMCs. PBMCs from two donors were infected with equal TCID50 of viruses generated from CAP137: two viruses representing A at 0 mpi, one representing B at ~2 mpi and two viruses representing AB viral population at 12 mpi. Viral replication was quantified by p24 ELISA at day 7, 10, and 14 post-infection. A) The mean p24 concentration from the two independent assays from 2 donors were plotted over time. Error bars represent SD. HIV-1 NL4-3 virus was used as positive control and cells only was used as negative control. Viruses A, B and AB are indicated with red, blue, and purple, respectively, and clone names are indicated in the key. B) Mean slopes of p24 curves were calculated using the slopes between RLU values at days 0 and 7, 0 and 10, and 0 and 14 (Weber et al., 2011). All slope values for each virus was used to calculate the mean and SD of two PBMC donors. The mean slope of all viruses was compared to virus A (137c1) using One-way Anova test with Bonferroni's multiple comparison test. $P < 0.05$ (*), $p < 0.01$ (**), $p < 0.001$ (***), $p < 0.0001$ (****).

CAP267

Four replicative IMCs were generated, two representing A and B viral populations at 0 mpi and two representing A and B viruses at 12 mpi. Compared to CAP137 IMCs, all the clones had lower replication capacity than NL4.3 suggesting that they were less fit variants. The slope of 267c6 replication kinetics was 6-fold higher than 267c2 ($p < 0.01$), suggesting that virus B variants gained replication fitness over time. However, the slope of 267c5 at 12 mpi

was 3-fold significantly lower than 267c1 at 0 mpi ($p < 0.01$) suggesting that virus A lost fitness over 12 months infection (Figure 4.4A and B).

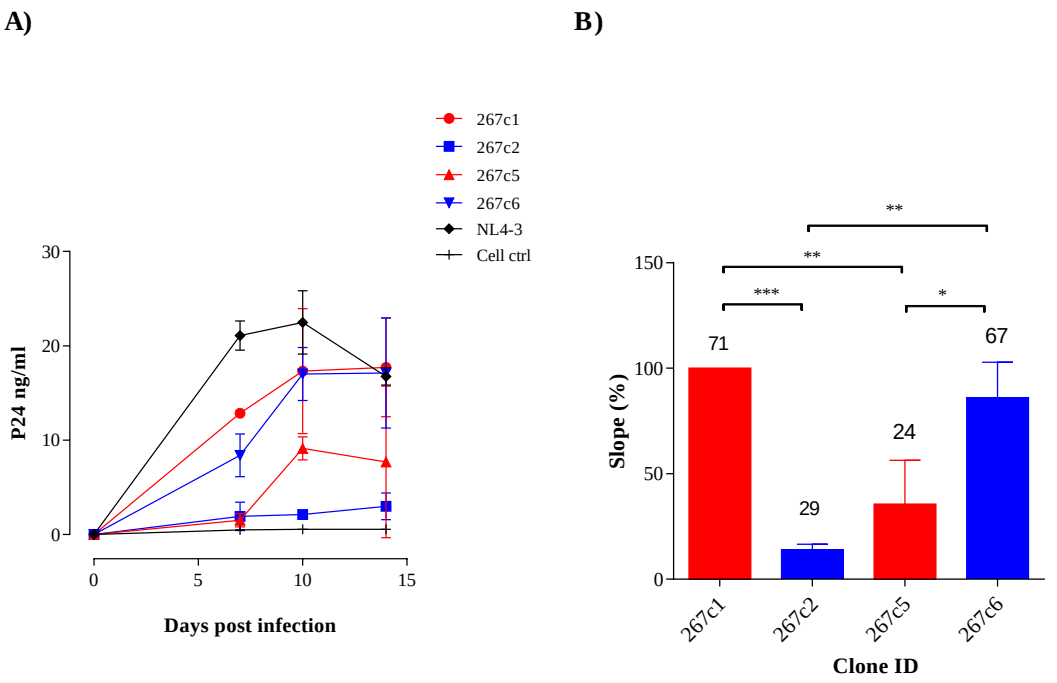


Figure 4.4. Replication of CAP267 chimeric IMCs. IMCs generated from CAP267, representing A and B viruses at enrolment, and at 12 mpi were compared in parallel using a PBMC replication assay. A) The mean p24 concentration of two independent experiments using two donors are indicated. P24 ELISA was measured from viral supernatant collected at day 7, 10, and 14 post-infection. Error bars representing SD. B) Mean slopes were calculated using the slopes between RLU values at days 0 and 7, 0 and 10, and 0 and 14 (Weber et al., 2011). The mean and SD of the three slope values were determined per donor and then the average was determined for the two donors. A and B viral populations are indicated in red and blue, respectively. Clone ID is indicated in the key. One-way ANOVA was used to compare the mean difference of slope between viruses, and the significant difference between viruses was calculated using Bonferroni's multiple comparison test indicated with the asterisk: * = $p < 0.05$ (*), $p < 0.01$ (**), $p < 0.001$ (***), $p < 0.0001$ (****).

4.4.4 Association between viral replication in PBMCs and *in vivo* frequency of variants

This study hypothesized that viral outgrowth in dual infected individuals was due to enhanced Env entry efficiency. There was no consistent relationship between *in vivo* outgrowth of variants and the PSV entry efficiency of all clones at each time point. However, recombinants that outgrew other variants infecting CAP37, CAP84 and CAP137 had enhanced entry efficiency by 12 mpi (Figure 2.10, Chapter 2), suggesting that at 12 mpi, Env entry could be a determining factor of *in vivo* viral fitness. To confirm whether this relationship was upheld within the context of multiple rounds of replication, we evaluated the relationship between *in vivo* frequency of variants and chimeric IMC replication of CAP137 and CAP267.

CAP137

We were unable to conclude whether virus A clones, 137c1 and 137c2, had similar RC at 0 mpi but given the experimental variation at 10 days post-infection, it is likely that 137c1 had higher RC than 137c2. The recombinant viral population at 12 mpi had the highest average RC and had outgrown all other variants at that time point (Figure 4.5A). Closer inspection of the AB population indicated that AB/B with 2-fold higher frequency had significantly lower RC than 137c9, the AB/A variant (Figure 4.3). This suggests that factors other than viral RC might have influenced the *in vivo* outgrowth of virus AB/B.

CAP267

The RC of CAP267 IMCs was associated with changes in *in vivo* frequency of corresponding viral populations. Virus A RC decreased ~3-fold over time which correlated with the ~2.4-fold drop in *in vivo* frequency of this viral population at 12 mpi (comprised 24 % of all variants at 12 mpi) (Table 2.1). Virus B, which dominated with a frequency of 76 % at 12 mpi, had significantly higher RC than virus B at 0 mpi (Figure 4.4), suggesting that viral population B evolved to enhanced replication fitness which then contributed to *in vivo* outgrowth (Figure 4.5B).

Despite instances when CAP137 IMC RC did not correlate with changes in *in vivo* viral outgrowth, the two participants, CAP137 and CAP267, were infected with viral populations

with enhanced replication fitness and high *in vivo* frequency at 12 mpi. This suggests that RC was enabling the recombinants to outcompete other variants at 12 mpi (Figure 4.5).

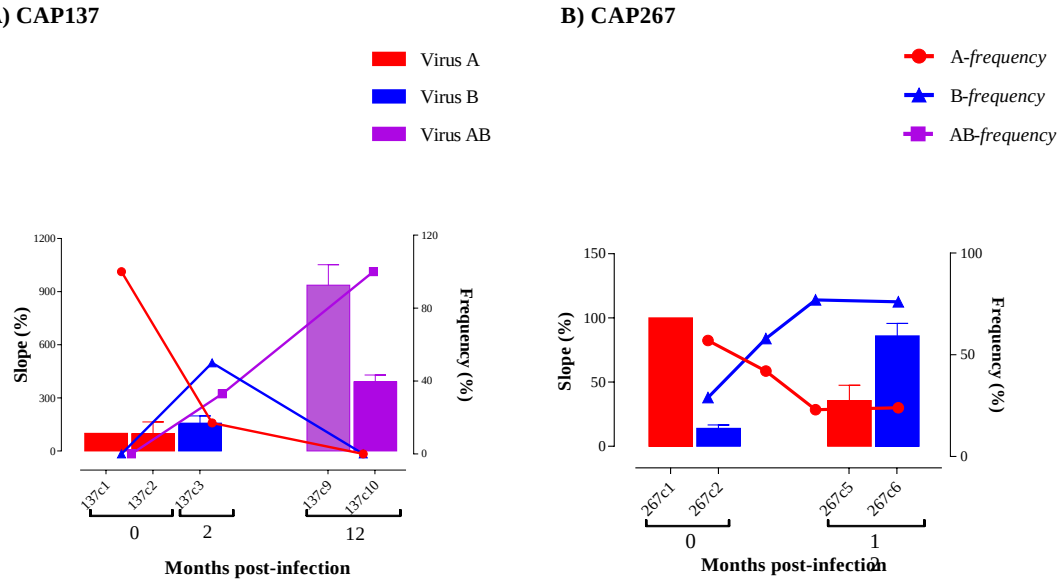


Figure 4.5. Association between Replication Capacity and in vivo frequency of viral populations over time.
Association between replication capacity (RC) of chimeric IMCs and in vivo frequency of viral populations over time showing an increase in IMC RC and population frequency at 12 mpi in A) CAP137 and B) CAP267. In vivo frequency (%) of variants was determined by calculating the relative proportion of sequences at each time point for each participant (Table 2.1, Chapter 2). Only frequency of CAP137 variants at 0, 2 and 12 mpi is shown. Virus A, B, and AB frequency is indicated by red, blue and purple lines, respectively. RC was expressed as the mean slope of curves of two independent donors and was compared to virus A over time. RC of virus A, B and AB is indicated by red, blue and purple bars, respectively.

4.4.5 Association between viral replication in PBMCs and PSV Entry Efficiency

We have shown that overall changes in PSV entry efficiency and chimeric IMC RC over 12 months of infection tracked the *in vivo* frequency of their source viral population, suggesting that Env was playing an important role in viral outgrowth *in vivo*. We next determined whether Env entry efficiency influenced the RC of matched chimeric IMCs for CAP137 and CAP267. Viruses at 0 mpi and 12 mpi were compared as PSV entry efficiency seemed to

fluctuate during the earlier time points and were not always linked to changes in population frequency.

CAP137

When we compared chimeric IMC RC, 137c9 RC was significantly higher than that of 137c10 (Figure 4.3), whereas the opposite was true when comparing PSV entry efficiency of matched Env clones (Figure 2.9C, Chapter 2). In Chapter 2, we suggested that AB/B virus might carry potential CTL escape mutations that influenced differences in PSV entry efficiency but here we show that AB/B IMC RC is significantly lower than 137c9. However, when the average entry efficiency of clones representing virus A at 0 mpi and AB recombinants at 12 mpi were compared, there was a clear association between the enhanced PSV entry efficiency and chimeric IMC RC, at the population level, at 12 mpi, suggesting that Env is an important determinant of viral fitness for CAP137 (Figure 4.6A and C).

CAP267

Overall, RC differences between IMCs followed the same trend as PSV entry efficiency suggesting that Env entry efficiency was driving RC resulting in outgrowth of variants (Figure 4.4A and B): 1) virus A (267c1) replication kinetics was 7-fold higher than that of virus B (267c2) at 0 mpi ($p < 0.01$); 2) replication kinetics of virus B 267c6 was 2.4-fold higher than virus A (267c5) at 12 mpi ($p < 0.5$); 3) 267c2 had the lowest RC at 0 mpi and 267c6 had the highest at 12 mpi (Figure 4.4). The RC of 267c6 was 6-fold higher than that of 267c2 ($p < 0.01$) which was almost identical to the difference in PSV entry. However, unlike PSV entry efficiency where virus A at 12 mpi was similar to virus A at 0 mpi, the slope of 267c5 at 12 mpi was ~3-fold significantly lower than 267c1 at 0 mpi ($p < 0.01$). The difference observed between 267c1 PSV entry and IMC RC is likely due to experimental variation**.**

Despite the discrepancies between individual variants, there was a significant correlation ($p = 0.03$, $r = 0.7$) between the RC of chimeric IMCs in PBMCs and their corresponding PSV entry efficiency in TZM-bl cells (Figure 4.7A). Similar results were observed when correlation analysis was performed using the mean AUC to represent IMC RC (data not shown). However, when IMC RC of Env isolated from CAP137 and CAP267 was compared

to their corresponding PSV entry efficiency in Affinofile cells that mimic T-lymphocyte ($CD4^{high}/CCR5^{low}$), no significant association was observed ($p = 0.3$, $r = 0.4$) (Figure 4.7B). This finding was contrary to our previous data which showed a significant association between PSV entry efficiency in TZM-bl cells and entry of Affinofile cells ($CD4^{high}/CCR5^{low}$) (Figure 3.6, Chapter 3). The absence of correlation could be due to a number of factors such as the influence of cell type (Joseph et al., 2014) combined with differences between IMC replication and PSV entry assays (Etemad et al., 2009; Murakami and Freed, 2000b). Furthermore, comparison between IMC RC and Affinofile PSV entry was limited to variants of CAP137 and CAP267 previously suggested to being sensitive to changes in cell type and HIV-1 backbone (Table 3.3). Sample size also has a large impact on correlation statistics and thus the loss of a few points could skew the outcome of the analysis (Vaux, 2012).

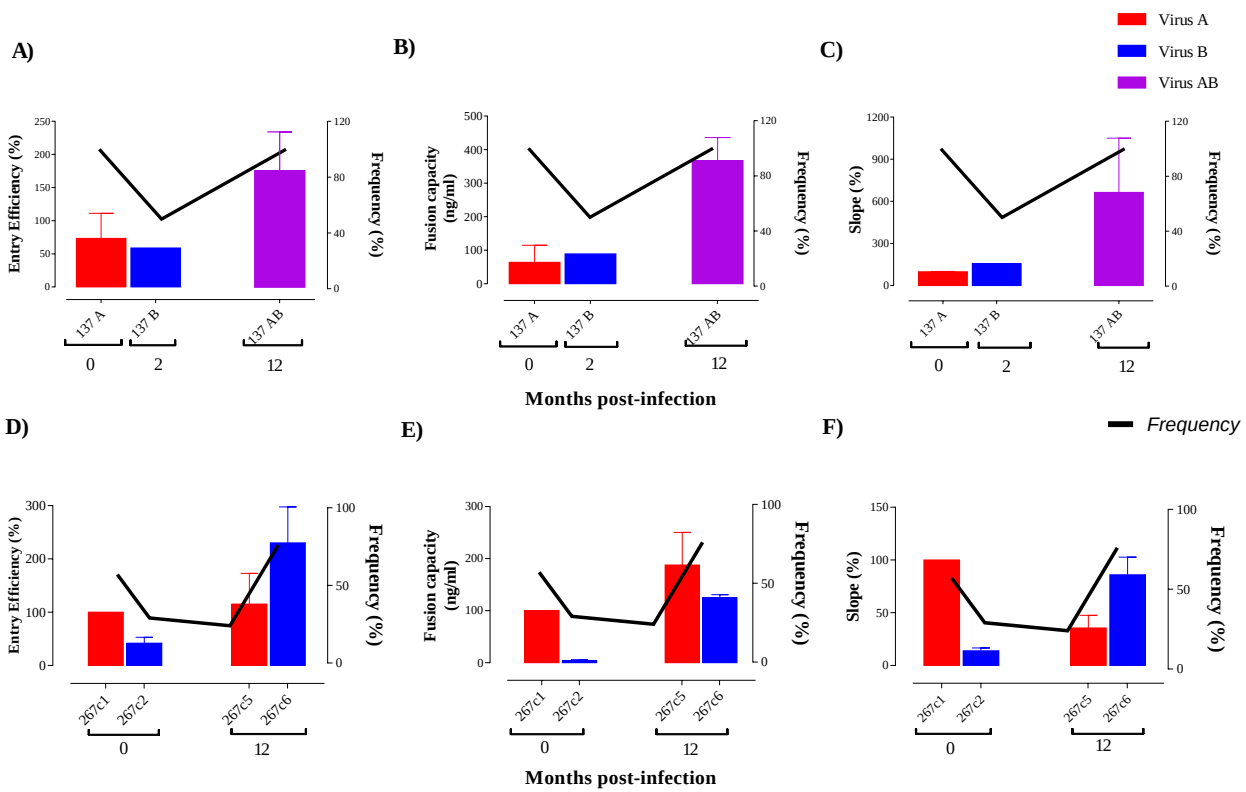


Figure 4.6. Association between overall outgrowth of viral populations and Env function. Association between the *in vivo* frequency of dominant viral populations of CAP137 with A) PSV entry efficiency; B) Fusion capacity; and C) Replication capacity. The increase in the *in vivo* frequency of CAP267 viral B populations at 12 mpi was associated with D) PSV entry efficiency; E) Fusion capacity; and F) Replication capacity. In cases where more than one clone represented a viral population at a particular time point then the average of entry efficiency, fusion or RC of the clones was calculated and used to represent the entry, fusion or RC of the viral population at that time point. Viral population A, B and AB are indicated in red, blue and purple, respectively. *In vivo* frequency (%) of variants was determined by calculating the relative proportion of sequences at each time point for each participant (Table 2.1, Chapter 2) and the frequency of only dominant viral population at each time point was shown by black line.

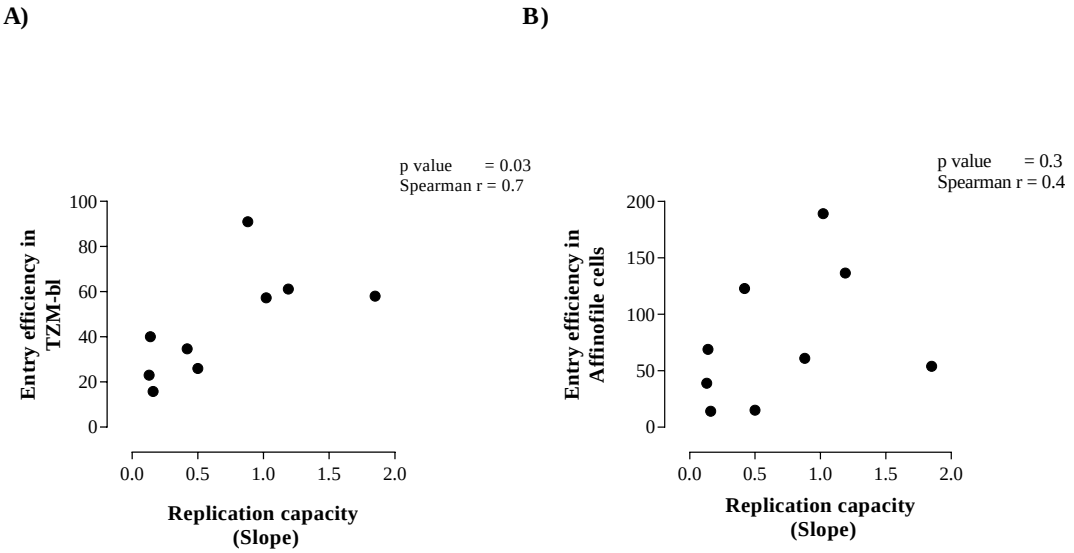


Figure 4.7. Correlation between IMC Replication capacity and PSV entry efficiency. Spearman r correlation was carried out to determine the relationship between viral replication capacity (RC) in PBMCs and PSV entry efficiency in A) TZM-bl and B) Affinofile cells that mimic T-lymphocytes. RC is expressed as the mean slope of replication kinetics of two independent experiments and PSV entry efficiency was normalised to the cell only control. Spearman r test was determined as indicative of a negligible, weak, moderate, high and very high relationship based on previous reports (Hinkle et al., 2003), p value of < 0.05 was considered significant. GraphPad Prism 5.0 software (CA, USA) was used for statistical analysis.

4.4.6 Association between viral replication in PBMCs and Env fusion capacity

In chapters 2 and 3, we suggested that gp41 might play an important role in determining PSV entry efficiency by influencing Env fusogenicity. As there was an apparent association between TZM-bl PSV entry and chimeric IMC RC, we determined whether IMC RC was also influenced by Env fusion capacity.

CAP137

Env fusogenicity followed a similar trend as IMC RC of matched variants across all time points. The difference between 137c1 and 137c2 IMC RC and PSV fusion capacity was

likely due to experimental error given the standard deviation of the independent experiments (Figure 4.3 and Figure 3.7). At 12 mpi 137c9 IMC had significantly higher RC than 137c10, similar to the difference in fusion capacity of the two clones except the relationship did not reach significance. When the average IMC RC and PSV fusogenicity of virus A and AB clones were compared instead of individual clones then RC and fusion capacity followed the same trend, suggesting that there is an association between Env fusogenicity and replication fitness of the viral populations infecting CAP137 (Figure 4.6B and C).

CAP267

The RC of 267c6 at 12 mpi was significantly greater than 267c2 and this corresponded to increased fusion capacity. However, whereas virus A IMC RC decreased over time, the fusion capacity of this viral population either increased ~2–fold or seemed to remain the same over 12 months of infection. (Figure 4.4; Figure 4.6E and F). This could suggest that fusion capacity contributed to the RC of virus B over time but not virus A. Although, virus A entry efficiency of TZM-bl and Affinofile cells, fusion capacity, and Env expression did not change over the course of infection, gp120 incorporation of 267c1 was higher than 267c5 (Figure 3.13), mimicking the difference in RC between the two clones. Although there was no correlation between gp120 incorporation into PSVs and PSV entry efficiency, perhaps the poor incorporation of 267c5 gp120 had a detrimental effect on IMC RC when measured over multiple rounds of infection. Further analysis is required to confirm the potential relationship.

Despite the variation between the replication fitness and fusion capacity of some clones, a significant correlation was determined between overall PSV fusion capacity and viral replication ($p = 0.04$, $r = 0.7$), suggesting that fusion capacity might play a role in *ex vivo* viral replication (Figure 4.8). A similar association was found between RC and fusion capacity when AUC was used as a measure of RC (data not shown).

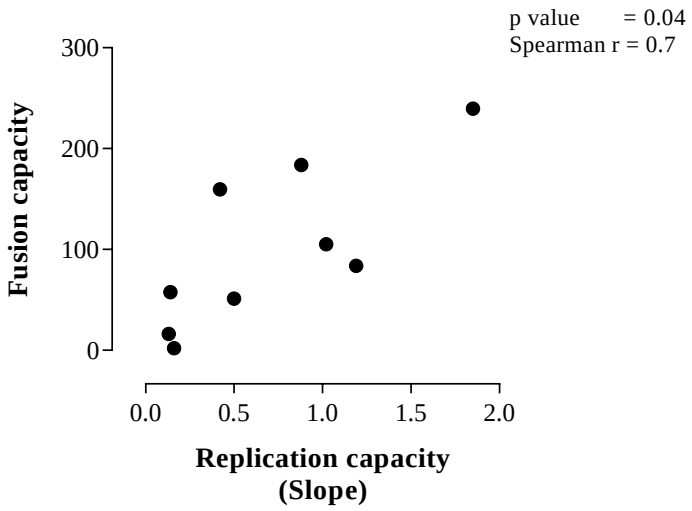


Figure 4.8. Correlation between RC and fusion capacity. Correlation analysis was carried out to determine the relationship between viral replication capacity (RC) and PSV fusion capacity. RC was expressed as the mean of the slope derived from replication kinetics. T-20 IC50 in (ng/ml) was used as an indicator of fusion capacity. Spearman r test was used to indicate a negligible, weak, moderate, high and very high relationship based on previous reports (Hinkle et al., 2003), p value of < 0.05 was considered significant. GraphPad Prism 5.0 software (CA, USA) was used for statistical analysis.

4.4.7 Association between viral replication in PBMCs and disease progression

HIV set-point viral load - the viral load (VL) at 12 mpi- is a predictor of disease progression (Mellors et al., 1996) and is driven largely by viral factors (Alizon et al., 2010; Hollingsworth et al., 2010; Mellors et al., 1996) suggesting that Env phenotype could influence how quickly HIV-infected individuals progress to AIDS. Our data supports this hypothesis as we showed that Env entry efficiency was significantly associated with CD4+ T cell loss (Figure 2.12, Chapter 2), a well-documented marker for disease progression (Phillips et al., 1991). As RC was previously shown to correlate with disease progression using both VL and CD4+ T cell count (Campbell et al., 2003; Quiñones-Mateu et al., 2000; Troyer et al., 2005), we determined whether the chimeric IMC RC of variants infecting dual-infected individuals revealed a similar association between RC and disease progression.

Increasing RC of chimeric IMC carrying CAP137 Env coincided with an increase in VL and sever decline in CD4 count over time with CD4+ T cells dropping to < 350 cells/µl in the first year of infection (Figure 4.9A). This suggested that the ability of variants to replicate to high titres due to Env entry efficiency within this individual might have contributed to her rapid disease progression. Similarly, the CD4 count of CAP267 also dropped to below 350 cells/µl within 1 year of infection and this decline tracked the RC of the dominant viruses over the course of infection (Figure 4.9B). However, VL decreased over 12 months of infection, suggesting that enhanced RC was not translating into high viremia. Due to small sample size, statistical analysis could not be performed.

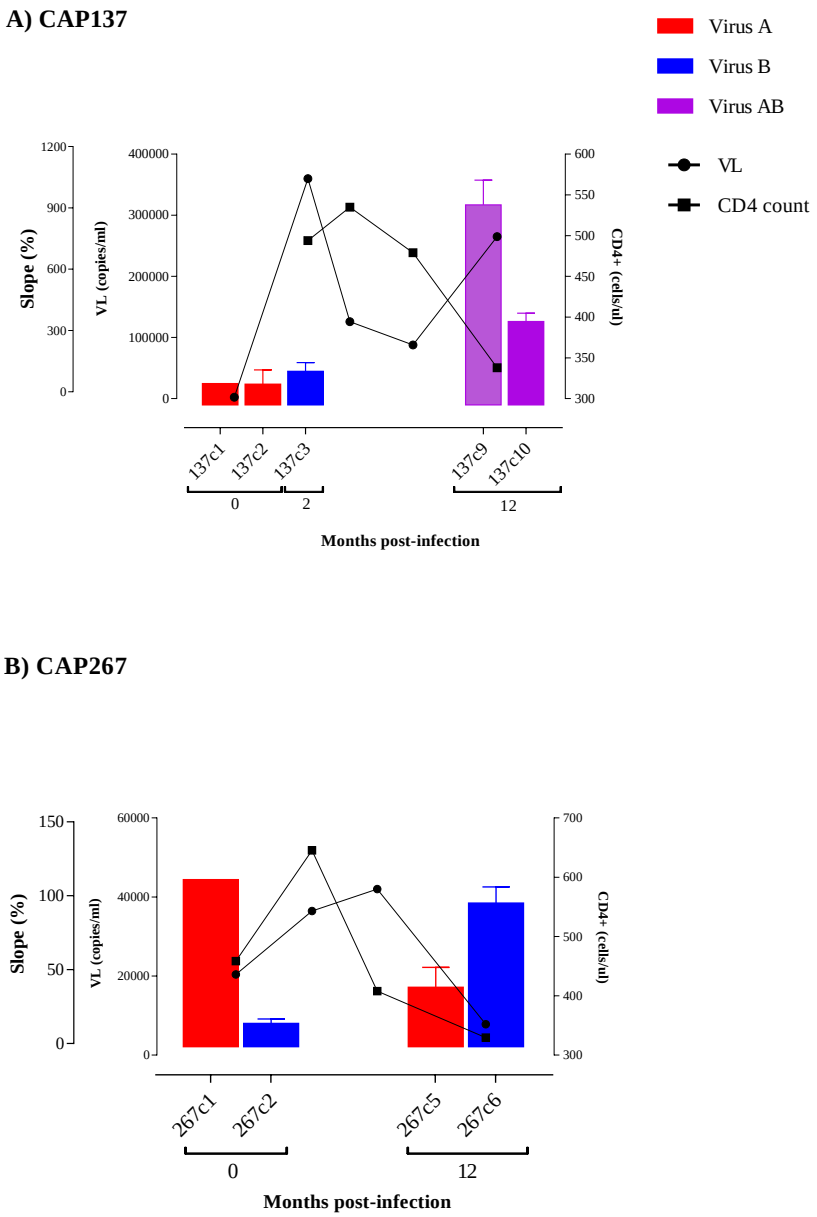


Figure 4.9. Association between replication capacity of chimeric infectious molecular clones and markers of disease progression.

Replication capacity (RC) of variants infecting A) CAP137 and B) CAP267 and markers of disease progression over time were plotted. Viral load (copies/ml) on the left y-axis and CD4 counts on the right inside axis vs months post-infection (mpi). The slope of the replication kinetics curve was used as an indicator of RC and indicated as a percentage (%) relative to virus A at 0 mpi to show the changes over time and plotted on the left outside y-axis. RC of A, B and AB variants are indicated in red, blue and purple bars, respectively. Changes in VL and CD4 count data over time are indicated by black circles and squares, respectively.

4.4.8 The impact of Env fitness determinants on IMC replication

In chapter 2, we identified gp41 and PNGs might play a role in Env entry and thus determine viral fitness. In order to determine whether these sites impacted the ability of the virus to replicate, we generated IMCs of the mutant clones of CAP137 and CAP267.

CAP137

As previously mentioned, a PNG at position 332 was deleted from clone 137c10 (AB viral population at 12 mpi) and the mutant, 137c10sdm had significantly lower entry efficiency than WT (Figure 2.14). When the replication of 137c10sdm was compared to WT, both viruses replicated to similar levels till day 10, after which 137c10sdm seemed to gain fitness, reaching a peak of 22 ng/ml p24 at day 14 (Figure 4.11A). However, there was no difference in either the slope or the AUC of 137c10sdm IMC RC compared to WT which differed from the 2-fold decrease in PSV entry efficiency of 137c10sdm. Furthermore, when the slope of 137c10sdm replication kinetics was compared to virus A at 0 mpi (137c1), the mutant had a 5-fold significantly higher RC ($p = 0.01$), whereas 137c10sdm had similar PSV entry efficiency to 137c1 (Figure 4.11B). This data suggested that the PNG at position 332, affected Env PSV entry efficiency without influencing the replication capacity of the virus.

CAP267

Virus B PSV entry efficiency was much higher at 12 mpi than at 0 mpi. We generated two chimeras where we swapped C5 and gp41 regions between the two infecting viruses (267c2/c6 and 267c6/c2) and found that the C5 and/or gp41 of the variant at 12 mpi was responsible for the increased entry efficiency. When the replication capacity of the chimeric viruses was compared to WT, the replication of both chimeras was significantly lower than 267c6 and equal to the replication level of virus B at 0 mpi (Figure 4.12). This result did not correlate with PSV entry efficiency suggesting that C5 and/or gp41 of 267c6 was not influencing IMC replication.

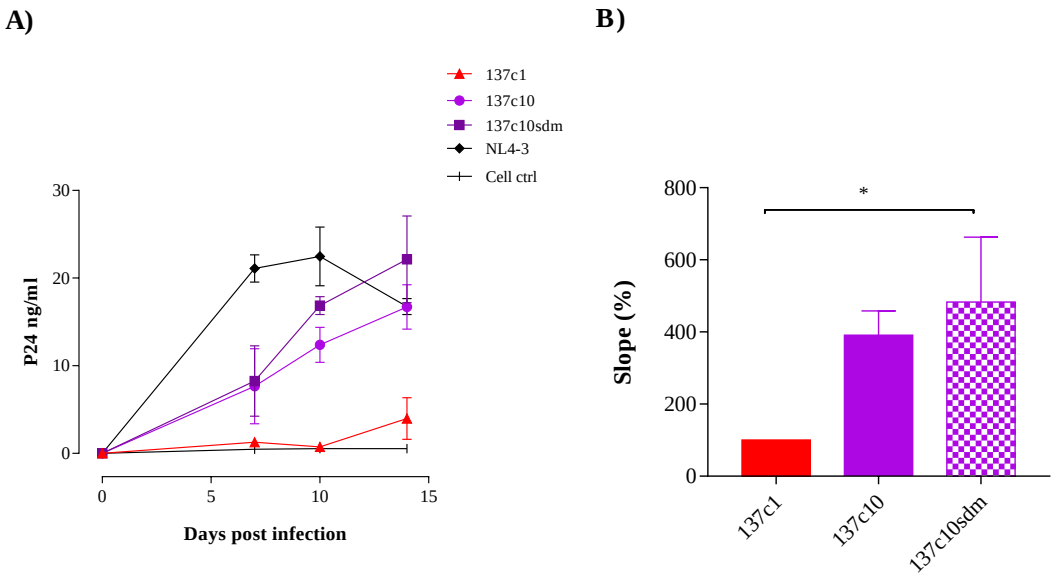


Figure 4.10. Impact of CAP137 PNG at position 332 on infectious molecular clone replication capacity. Viruses representing viral population A at enrolment (137c1), AB virus at 12 mpi (137c10) and the AB mutant (137c10sdm) were compared in two independent replication assays using PBMCs. A) The mean p24 concentration of the two independent assays were plotted over days post-infection and error bars represent standard deviation. HIV-1 NL4-3 virus was used as a positive control and cells only was used as negative control. Viruses A, AB and AB-sdm are indicated in red, light purple, and dark purple, respectively, and clone names are indicated in the key. B) Mean slope of each curve from independent experiments were calculated using the slopes between days 0 and 7, 0 and 10, 0 and 14. The overall slope of each curve from the two independent biological repeats was averaged and represented as bars with error bars representing SD. The mean slope of 137c10 and 137c10sdm was compared to virus A (137c1) using One-way Anova test with Bonferroni's multiple comparison test. $p < 0.05$ (*), $p < 0.01$ (**), $p < 0.001$ (***), $p < 0.0001$ (****).

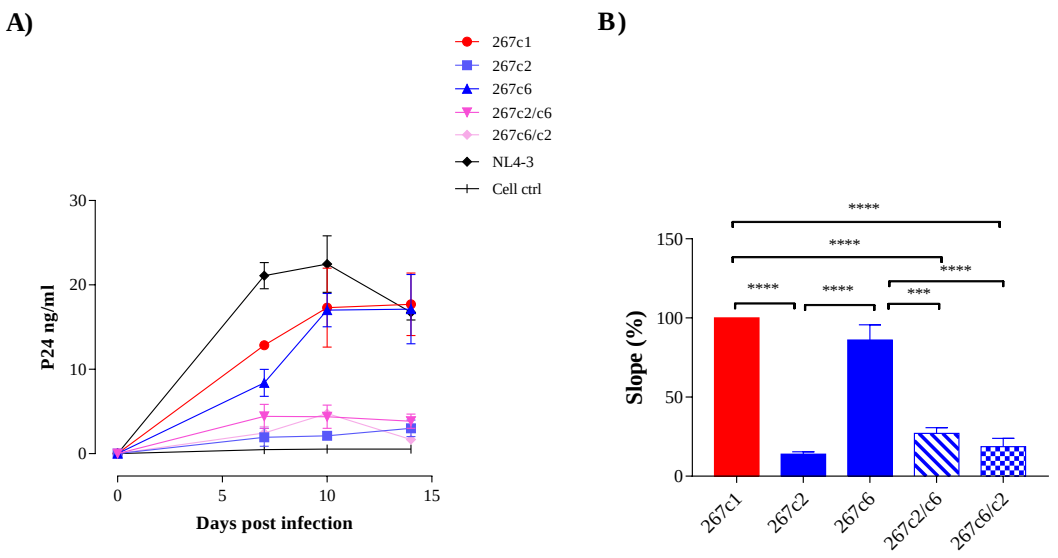


Figure 4.11. Replication Capacity of CAP267 chimeras. To test the importance of gp41 in PSV entry efficiency, restriction enzyme (RE) digestion of 267c2 and 267c6 was used to swap the C-terminus between the two Envs using corresponding RE sites. RE digestion swapped gp41 and C5 between the two clones but as C5 only differed by one amino acid residues between the two chimeras, gp41 is likely to be the important determinant of PSV entry efficiency. The two chimeras of CAP267 B viruses were compared with WT viruses in the replication assay to determine whether gp41 was important to viral RC. A) The mean p24 concentration change over days infection of two independent experiments using two donors are indicated. P24 ELISA was measured from viral supernatant collected at day 7, 10, and 14 post-infection. Error bars represent SD of two independent experiments. B) The overall slope of each p24 curve generated per independent experiment was calculated by finding the mean of the slopes between days 0 and 7, 0 and 10, and 14. The overall slope of each curve was then averaged for each virus and standard deviation of two independent biological repeats is indicated. Viral population A, B are indicated in red and blue, respectively. The Chimeras 267c2/c6 and 267c6/c2 are indicated in blue striped bar and blue blocks, respectively. One-way Anova was used to compare the mean difference of slope of the viruses, and the significant difference between viruses was calculated using Bonferroni's multiple comparison test indicated with the asterisk: * = $p < 0.05$ (*), $p < 0.01$ (**), $p < 0.001$ (***), $p < 0.0001$ (****).

4.4.9 Summary of results

Table 4.1. Ranking of replication capacity of IMCs in PBMCs

Participant ID	IMC Replication capacity in PBMCs	Pseudovirion entry efficiency in TZM-bl cells	Env clones that differed between assays
CAP137	c9>c10>c1≈c2≈c3	c10>c9≈c1>c2≈c3	c9, *c10, c1
CAP267	c1≈c6>c5>c2	c6>c5≈c1>c2	*c1

*137c10 and 267c1 was suggested to be sensitive to changes in cell-type and/or HIV-1 backbone differences (Table 3.3, Chapter 3).

Table 4.2. Summary of associations between infectious molecular clone replication capacity and *in vivo* outgrowth and other Env phenotypes

Association of Increased replication capacity with variant outgrowth		Significant association between increased replication capacity, viral outgrowth and Envelope phenotypes					
Over 12 mpi	Within time points	Entry efficiency in TZM-bl cells	Fusion capacity	Expression	Gp160 cleavage	CD4 count	VL
Yes	*Yes	*Yes	*Yes	No	No	*Yes	No

*Association was not apparent for all clones

Table 4.3 Impact of Env fitness determinants on chimeric infectious molecular clone replication capacity

Participant ID	Fitness determinant of pseudovirion entry efficiency	Mutants generated	Association with IMC replication capacity
CAP137	PNG at N332	Deletion of N332 in 137c10	No
CAP267	5 amino acids in cytoplasmic tail of gp41 and one amino acid in C5 (gp120)	chimeric Env clones (267c2/c6 and 267c6/c2)	No

4.5 Discussion

This study characterized the RC of chimeric IMCs that carry Env clones from variants infecting dual infected individuals over the first year of infection. Two participants classified as rapid progressors (CAP137 and CAP267) were selected for this study as both were infected with variants at 12 mpi with high entry efficiency with concomitant increased *in vivo* frequency but differed in the presence of recombinants. We succeeded in generating 9 IMCs that replicated within PBMCs generated from two donors in independent experiments. The low number of functional clones could be due to the incompatibility between subtype C *env* and the subtype B NL4-3 backbone (Bagaya et al., 2015).

RC of viruses using parallel replication assays over 2 weeks was measured using the area under the p24 curve and (Charpentier et al., 2006) the slope of the curve (Weber et al., 2011). Despite differences in the degree of significance, both methods indicated the same relationship between IMC RC and Env phenotypes and we thus only reported on measurements using the slope of the replication kinetics curve.

We compared frequency of variants and changes in IMC RC of viral populations to determine whether RC could be influencing *in vivo* variant outgrowth. RC of chimeric IMCs representing the recombinant population infecting CAP137 had out-competed viruses A and B at 12 mpi and had the highest RC of all variants, suggesting that recombination within Env had selected for variants with increased RC that enabled outgrowth. Virus B variants infecting CAP267 gained RC over time with corresponding increased frequency whereas virus A RC decreased with concomitant lowered frequency. Furthermore, at 0 mpi and 12 mpi the variant with the highest frequency also had the highest RC, suggesting that virus B had evolved within Env, conferring a competitive advantage. The ranking of CAP267 and CAP137 IMC RC from high to low was similar to the relative PSV entry efficiency of the clones (Table 4.1; Table 3.3) and there was an overall significant correlation with PSV entry efficiency ($p = 0.03$, $r = 0.7$), suggesting that Env fitness was a determinant of viral outgrowth. Therefore, at the population level, not only did PSV entry efficiency correspond to variant frequency but also RC, supporting the positive relationship between viral outgrowth and enhanced Env function (Figure 4.6).

However, when individual variants infecting CAP137 were compared, virus AB/A at 12 mpi had higher RC than AB/B despite its lower *in vivo* frequency. In addition, the ranking of 137c1, 137c9, 137c10 and 267c1 IMC RC and PSV entry efficiency from high to low did not correlate (Table 4.1). There were also instances where PSV entry efficiency did not follow variant frequency, suggesting that variants might not always reflect population changes. The reasons why individual clones did not reflect population dynamics could be multifactorial such as the impact of immune escape, HIV diversity and the number of clones analysed. As the sample size of our study was small, with only nine functional IMCs, we might not have sampled sufficient variants to observe population level RC. However, similar to PSV entry efficiency, population-based RC correlated significant with PSV fusion capacity ($p = 0.04$, $r = 0.7$) suggesting that fusion capacity was influencing virus RC (Table 4.2). Furthermore, similar to PSV entry, population RC was associated with CD4+ T-cell decline suggesting that high RC could hasten disease progression (Claiborne et al., 2015; Garg and Joshi, 2017). Therefore, RC could play a role in the rapid disease progression of CAP137 and CAP267 (Campbell et al., 2003; Lassen et al., 2009; Quiñones-Mateu et al., 2000; Trkola et al., 2003; Troyer et al., 2005).

We previously suggested that gp41 was a potential fitness determinant of variants infecting CAP137 and CAP267 (Table 4.3). As fusion capacity of these variants was apparently associated with IMC RC, it seemed probable that recombination within gp41 had resulted in enhanced RC of viruses infecting CAP137 and CAP267. However, contrary to the PSV entry efficiency and fusion capacity of 267c2/c6 that remained similar to 267c6, the RC of 267c2/c6 decreased significantly, indicating that the gp41 of 267c6 was not responsible for the low RC of this IMC. Therefore, even though 267c6 gp41 seemed to enhance PSV entry efficiency and fusogenicity, it did not impact IMC RC. Similarly, PNG 332 that seemed to play a significant role in the enhanced PSV entry efficiency of 137c10 had no effect on the RC of the AB/B recombinant. These unexpected results suggested that mutations can affect Env entry efficiency and fusogenicity without affecting the ability of the virus to replicate, despite the apparent association between PSV entry efficiency and IMC RC of the WT variants. However, the influence of PNG332 on Env function might be highly sensitive to changes at other a.a. residues as it had very different effects on YU2 and JR-FL (Lavine et al., 2012).

As the single-round PSV infection assay measures differences in entry, the first step in HIV replication, any differences detected should also be apparent in IMC infection assays which measures multiple replication cycles. It is thus likely that the inconsistent association between PSV entry efficiency and IMC RC between the mutants and WT viruses is due to differences in the assays. Provine *et al*., (2009) reported that IMC phenotype did not always resemble that of matched PSVs (Provine et al., 2009) potentially due to cell-specific differences between HEK293T cells PSVs and PBMCs (Herrera et al., 2005; Murakami and Freed, 2000b; Provine et al., 2009; Pugach et al., 2007). It was previously suggested that factors that interact with Env vary between cell-types so that mutations that influence Env entry efficiency in one cell type might have no effect in another (Murakami and Freed, 2000b). As replication in PBMCs represent a more physiologically relevant system to identify viral factors that could modulate HIV infection, it is possible that gp41 of 267c6 and N332 of 137c10 are not fitness determinants *in vivo*. However, as RC was determined using PBMCs from only two donors, additional experiments could prove otherwise.

Conclusion

In general, high IMC RC representing viral populations A, B and AB was associated with increased PSV entry efficiency, PSV fusogenicity and *in vivo* frequency, suggesting that Env played an important role in viral fitness and *in vivo* outgrowth. Although, the virus at the highest frequency at any given time would most likely be driving viremia and thus disease progression, IMC RC did not correlate with viral load. However, IMC RC was associated with CD4 T cell decline suggesting immune responses could be limiting viral load but not cytopathogencity. Therefore, RC could be influencing disease progression. However, we were unable to confirm the importance of gp41 and PNG332 in PSV entry efficiency as RC of matched IMC chimeras and mutant were not similarly affected. This highlights the importance of differences between assay systems and the need to standardise methodologies when identifying potential fitness determinants for vaccine and drug targets.

CHAPTER 5.

Summary and Conclusion

5.1 Discussion

HIV-1 vaccines have a small window of opportunity to elicit an immune response that could effectively prevent colonisation of mucosal lymphoid tissues and establishment of viral latency. One way to prevent HIV infection would be to stimulate broadly neutralising antibody (bnAb) responses that target Env, the viral component that mediates viral entry of CD4+ T cells. However, there are many challenges facing the design of an efficacious vaccine such as high viral diversity within and between individuals, the lack of a relevant animal model, evasion of immune responses by numerous viral mechanisms such as MHC down-regulation and the sequestering of CD4 and CCR5 receptor binding sites within inaccessible regions of Env. Vaccine design focussed on eliciting bnAb that prevent Env-mediated viral entry include numerous strategies including the use of purified recombinant gp120/gp140 monomers, Env trimers, hyperglycosylated Env and Env covalently bound to CD4 to induce exposure of receptor binding sites. The extensive diversification of Env however impedes the generation of neutralising antibodies that prevent infection of a broad range of viruses (Girard et al., 2006).

It has been suggested that conserved sites essential for the maintenance of Env structure and function, and more accessible than CD4 and CCR5 receptor binding sites, might make ideal targets for vaccine design. Recently, deep mutational scanning has indicated that epitopes of bnAbs are highly sensitive to mutation, suggesting that the identification of Env fitness determinants sensitive to change might be an alternative to apparent random selection of Env subunits as vaccine candidates (Haddox et al., 2016). One hindrance to the identification of fitness determinants is the flexibility of Env structure to accommodate multiple sequence changes without losing function. Highly variable host immune responses coupled with extensive viral diversity will alter rates of evolution and generate highly heterogeneous variants with varied functional stability. However, one can assume that even a structure as plastic as Env has a preferred energy state most suited for its function. Directed evolution selected for variants that differed in sequence changes

but had homogeneous functional stability and when these mutations were combined in a single variant, it became the most stable of all (Leaman and Zwick, 2013). This finding suggests that rapid HIV evolution within disparate hosts could select for fitter variants that ultimately share common phenotypes despite the introduction of different sequence changes. Furthermore, another study showed that functional constraints limited *in vitro* recombination between subtypes A and D to C1 region within gp120 and gp41 (Bagaya et al., 2015).

In dual infected individuals, phylogenetically distinct viruses compete for survival *in vivo* under the same host selective pressures. The co-circulation of two different strains could provide the sequence space for rapid selection of Env recombinants with high entry efficiency that speed up viral replicative fitness and disease progression (Quiñones-Mateu et al., 2000). Therefore, dual infections could provide the ideal model system to identify fitness determinants within conserved regions of Env that emerge quickly after infection. Dual infections could also be used to determine the importance of Env function in viral replication capacity (RC) and viral population frequency as well as identify the phenotype of Env most sensitive to genotypic changes. Therefore, this thesis hypothesised *that rapid viral evolution in dual infected individuals resulted in the early emergence of recombinants carrying Env fitness determinants that enhance viral replication capacity resulting in increased disease progression.* The frequency of dual infections varies between reports and is highly dependent on methods used to identify and quantify variants (Pacold et al., 2010). Our lab had previously identified only four dual infected individuals after screening 40 women- three were classified as rapid progressors and one a typical progressor- limiting the sample size and statistical power of the study. This thesis therefore aimed to describe the change in Env function over time and to determine whether changes tracked those of viral outgrowth to identify regions of Env under positive selection.

As the focus of the study was to characterise Env function, pseudovirus (PSV) was generated using Env cloned from longitudinal samples at 0, 3, 6 and 12 mpi. We hypothesized that the PSV single-round infection assay would be the best strategy to identify fitness determinants of Env. However, the limitation of this approach was that cell lines such as HEK293T cells were used to produce PSV and TZM-bl, U87 and Affinofile cells were used to monitor infection, instead of more physiologically relevant

cells. Furthermore, instead of using a subtype C backbone to determine PSV entry efficiency, two different subtype B vectors, pSG3 and pNL4-3.Luc.R^{-}E^{-} backbones lacking *env* were used to monitor infection of the different cell lines by luciferase reporter assays. However, the effect of different cell lines and HIV backbone vectors seemed to have minimum impact on PSV entry efficiency as most Env clones had the same relative entry efficiency irrespective of cell type and HIV backbone.

As changes in tropism have been linked to differences in Env function and disease progression (Koot, 1993; van Rij et al., 2003; Sucupira et al., 2012; Troyer et al., 2005) we firstly tested whether PSVs were able to enter cells with low CD4/high CCR5 density and cells with high CD4/high CXCR4 levels. All PSVs were unable to infect U87-CD4+CXCR4+ cells and able to infect Affinofile cells induced to express high CD4 and low CCR5 levels, similar to T lymphocytes, indicating that all variants were R5- and T-tropic. This finding supported previous studies that revealed no difference in viral tropism between transmitted and chronic viruses (Parrish et al., 2012; Ping et al., 2013).

A total of 295 SGA-derived amplicons were generated over time for all four participants and the proportion of sequences belonging to a viral population at each time point was used as a measure of relative outgrowth of variants. High diversity and early recombination events challenged the identification of distinct viral populations and it is highly likely that recombinants were present at the earliest time point for all participants but as both parent viruses were not detected we could not confirm this suggestion. Despite this limitation, the combination of a number of strategies- Neighbour-joining trees, Highlighter plots, RIP and pairwise DNA distance- identified two viral populations that grouped separately at the earliest time points allowing us to track their frequency over time and the emergence of recombinants.

Diversity was higher within the first 3-6 months and seemed to stabilize by 12 mpi suggesting that for all participants viral evolution was more rapid at earlier time-points. Abrahams et al (2013) suggested that immune responses drove rapid viral evolution during early infection until escape mutations became fixed within a population, lowering diversity (Abrahams et al., 2013). CAP84, classified as a typical progressor had lower sequence diversity than the other three participants, and it is interesting to speculate that slower evolution hindered the emergence of escape mutants, lowering viral load and slowing the disease progression of this individual.

We showed that evolution of the two viruses infecting 3/4 dual infected individuals resulted in the early emergence of recombinant viruses at 2 mpi for CAP137, 3 mpi for CAP84 and 6 mpi for CAP37. The emergent recombinants increased in *in vivo* frequency and outgrew other variants at 12 mpi. Previous studies showed an association between recombination and disease progression (Fang et al., 2004; Gordon et al., 2016; Nájera et al., 2002; Ramirez et al., 2008). However, in this cohort, the *in vivo* outgrowth at 12 mpi of emergent recombinants was associated with rapid disease progression for only two participants, suggesting that recombination is not always linked to disease progression (Table 5.1) (Liu et al., 1997). There was no apparent recombination between virus A and B for CAP267 although polymorphisms within the cytoplasmic tail (CT) of gp41 seemed to become fixed over time. The mapping of recombination events revealed that, although there was variation between participants, gp41 recombinants were apparently selected in 3/4 participants (Figure 2.7, Chapter 2), similar to previous reports (Bagaya et al., 2015). Potentially, these genotypic changes were responsible for enhanced PSV entry efficiency as recombinants at 12 mpi were able to infect TZM-bl cells better than other variants at the same time point. Therefore, at 12 mpi, 3/4 individuals were infected with recombinants that had out-competed other variants, most likely due to their average enhanced Env entry efficiency.

Although variants infecting CAP37 at 0 mpi had the highest level of diversity, emergent recombinants were detected later than those infecting the other participants. The high diversity made it difficult to identify regions that might have influenced PSV entry efficiency and viral outgrowth and thus the identification of fitness determinants focused on the other three participants (Table 5.2). By 12 mpi, CAP84 AB virus had evolved to include regions of virus B gp41 sequence as well as unique changes within gp120 and the signal peptide whereas the recombinants AB/A and AB/B infecting CAP137 had differences mainly within HR2 and MPER regions within gp41. Mutations were scattered across gp120 but most occurred in the CT of gp41 of the fit virus B infecting CAP267 at 12 mpi. As we wanted to determine whether survival pressure on the two variants infecting dual infected individuals selected for changes within a common region, we focussed on changes within gp41. Chimeras were generated for CAP276 and CAP84 where gp41 was swapped between two Env clones that either had different (CAP267) or

similar (CAP84) PSV entry efficiency associated with point mutations (CAP267) or recombination (CAP84) within gp41.

Escape from nAb and CTL immune responses have been shown to come with a fitness cost (Lynch et al., 2015; Moore et al., 2012; Wang et al., 2013). Chimeras of CAP84 and site-directed mutagenesis suggested that putative CTL escape mutations and the presence of a PNG at position 339 acted synergistically to decrease PSV entry efficiency and that recombination within gp41 restored fitness. The chimeras of CAP267 suggested that 5 amino acids within the CT (and/or 1 a.a. in C5) were involved in the increased fitness of the dominant virus at 12 mpi. Reversion of CTL escape mutation to consensus sequence has been shown to rescue any loss of fitness associated with immune escape (Song et al., 2014) Of the five a.a. in gp41 of 267c6, two reverted to consensus, suggesting that these two a.a. might play a role in CAP267 fitness. Further SDM might be able to confirm this relationship. CAP267 was HLA B*5802 positive but we could not detect any putative CTL mutations within the epitope associated with increased viral loads, QL11 (Kiepiela et al., 2007) nor HF9 identified by MotifScan (www.lanl.gov). CTL responses targeting Env have, in general, been associated with increased viremia, probably due to the plasticity of Env structure able to withstand amino acid changes (Kiepiela et al., 2004). However, without Elispot data, we cannot confirm the influence of CTL escape on disease progression and Env diversity of CAP84 and CAP267. Furthermore, PNG frequency did not increase consistently over time which suggested that bnAb pressure did not play a major role within the first year of infection as previously suggested (Abrahams et al., 2013).

The propensity of variants with high PSV entry efficiency and high frequency at 12 mpi to carry sequence changes within gp41 suggested that Env fusogenicity might play a role in influencing PSV entry efficiency as previously suggested (Brenchley et al., 2004; Gorry et al., 2014; Gorry and Ancuta, 2011; Lobritz et al., 2007; Marozsan et al., 2005; Rangel et al., 2003). There was an overall significant association between PSV entry efficiency and Env fusion capacity, with membrane fusion possibly the key Env function enhanced during the evolution of variants infecting CAP37 and CAP137. For CAP267 there was no obvious association between PSV entry efficiency and fusion capacity although there was an association between PSV entry efficiency and fusion capacity of CAP267 chimeras. This suggests that fusion capacity might be the driving factor

enhancing Env function, although this relationship did not hold true for some clones (Brenchley et al., 2004; Gorry et al., 2014; Gorry and Ancuta, 2011; Lobritz et al., 2007; Marozsan et al., 2005; Rangel et al., 2003).

On the other hand, differences in PSV entry efficiency of CAP84 wild-type (WT) clones and chimeras were unlikely due to Env fusogenicity. The apparent lack of association between PSV entry and Env fusogenicity for CAP84 chimeras suggested that other viral phenotypes might be influencing variant fitness in this participant. When Affinofile cells expressing 5- to 10-fold lower levels of CD4 and CCR5 than TZM-bl cells were infected with 84c1 and 84c4 PSVs, there was no significant change in entry of Affinofile cells over 12 mpi, similar to entry of TZM-bl cells but unlike PSV fusogenicity. CAP84 diversity seemed to be driven by escape from immune responses (Appendix B) and these polymorphisms might have selected for viral populations highly divergent in Env phenotype. As characterizing immune escape was outside the scope of this study, CAP84 phenotype was not further characterized.

As CAP37, CAP137 and CAP267 gp41 fusion capacity was apparently contributing to PSV entry efficiency and gp41 had been shown to influence Env cell-surface expression (Lebigot et al., 2001), we wanted to determine whether expression and processing of Env might be influencing Env fusion capacity. Due to the high diversity of CAP37, this participant was not included in this analysis.

Comparison between PSV entry efficiency and fusion capacity with cell associated Env levels and cleavage into gp120 and gp41, suggested that the more Env cleavage the higher PSV entry efficiency and fusogenicity although this relationship did not reach significance. We hypothesised that higher cell-surface expression of gp120 could increase the number of Env spikes incorporated into PSVs (Bachrach et al., 2005) and result in enhanced entry efficiency and fusogenicity (Chertova et al., 2002). However, there was no apparent relationship of Env expression and cleavage with incorporation of gp120 into PSVs. Possibly over-expression of Env might slow gp160 cleavage and folding during cell trafficking (Blay et al., 2007) that then differentially influences the incorporation of gp120 into PSVs. As incorporation of gp120 was also not associated with PSV entry efficiency, incorporation of functional trimers might not be limiting. Bachrach *et al* (2005) suggested that increased infectivity was only dependent on Env incorporation until a certain point after which it no longer had an effect (Bachrach et al.,

2005). Alternatively, robust entry of TZM-bl cells due to high levels of CD4 and CCR5 could mask subtle influences that Env expression, cleavage and incorporation might have on PSV entry. If this was the case, then TZM-bl cells might not be good model for HIV infection of T lymphocytes. However, there was a significant association of PSV entry of TZM-bl cells with entry of Affinofile cells induced to express high CD4/low CCR5 levels and with the RC of matched chimeric IMCs, suggesting that PSV infection of TZM-bl cell model system was physiologically relevant.

When comparing TZM-bl entry with other Env phenotypes, such as fusion capacity, p values less than 0.05 was used to confirm whether one phenotype was associated with another. However, in some cases differences between entry efficiency, and fusogenicity of clones did not reach significance despite apparent trends. In addition, differences between independent experiments due to intra-experimental and inter-experimental variation made it difficult to confirm associations. To support the power of our p values (Nuzzo, 2014) given the small sample size, we applied multiple mechanistic experiments to confirm our findings in PSV entry of TZM-bl cells such as infection of different cells with varying receptor levels and alternative HIV backbones (Vaux, 2012). Lastly, we also compared the RC of chimeric IMCs to confirm the importance of Env differences. Furthermore, after identification of putative fitness determinants by sequence analysis we hypothesised that mutagenesis of Env clones could also be used to confirm relationships between TZM-bl entry and Env phenotypes.

Similar to PSV infection of TZM-bl cells, variants infecting CAP137 and CAP267 had increased RC at 12 mpi and this coincided with high *in vivo* frequency of viral population. This suggested that PSV entry of TZM-bl cells represented changes in RC and facilitated the outgrowth of variants at 12 mpi. A similar relationship was observed between IMC RC and PSV fusion capacity, supporting our suggestion that Env fusogenicity could be responsible for differences in PSV entry efficiency. However, when we compared PSV entry and fusogenicity with IMC RC of Env mutants, a similar relationship was not observed (Table 5.2). The IMC replication capacity of CAP267 chimeras was lower and IMC RC of CAP137c10sdm lacking a PNG at N332 did not change compared to WT, in conflict with the decreased PSV entry efficiency of the mutants compared to the dominant variants at 12 mpi. Therefore, although IMC RC supported changes in PSV entry efficiency and fusogenicity of WT clones, it did not confirm the influence of gp41 and

N332 on *in vitro* viral replication in CAP267 and CAP137, respectively. Therefore, despite previous studies showing that gp41 played an important role in Env entry efficiency (Baatz et al., 2011; Lobritz et al., 2007; Marozsan et al., 2005; Reeves et al., 2005) this did not translate into IMC replication fitness. There are differences between pseudovirion and IMC infectivity assays that might have contributed to the differences between the two study outcomes (Provine et al., 2009). Furthermore, differences between *in vitro* IMC mono-infections and PSV entry efficiency could be due to the detection threshold of the replication assay. Although, competitive replication assays are more sensitive and might better represent *in vivo* viral fitness they are limited by *in vitro* recombination and was thus not considered in this study (Quiñones-Mateu and Arts, 2001). Alternatively, perhaps determinants of replicative fitness should be identified first and then tested for a role in Env functions using PSV assays. We would also like to suggest that the effect of mutations on Env function and/or processing is highly clone specific due to amino acid interactions between different regions of Env: a salt-bridge required for CD4 binding in one clone might not be essential in another. Our lab has discovered that the introduction of PNGs can abrogate function of one clone and leave another unaffected (Manuscript in preparation).

Two PNGs at positions 332 and 339 were identified that increased and attenuated Env fitness, respectively, possibly because gain or loss of N-glycans that facilitated escape from immune pressure came with a fitness cost (Cong et al., 2007; Liu et al., 2007; Song et al., 2012; Sunshine et al., 2015). PNGs play a vital role in immune escape and neutralizing antibody recognition (Moore et al., 2012) and influence Env fitness (Lavine et al., 2012; Wang et al., 2013). The effect of N332 on CAP137 Env fitness is noteworthy given that this N-glycan is responsible for eliciting bnAb and might make a good target in vaccine design. However, if it alters Env fitness then loss or gain of the PNG due to immune pressure could potentially generate fit variants able to out-compete others and drive viral loads and disease progression, mitigating the positive effect of the vaccine. This highlights the importance of screening potential vaccine targets for influencing viral fitness.

Three of the four participants in this study were classified as rapid disease progressors and there was an association between PSV entry efficiency and CD4 decline (Table 5.1). This relationship was upheld when RC was compared to CD4 counts although sample

size precluded statistical analysis. This, suggested that both Env function and viral RC could influence disease progression, supporting previous studies (Gordon et al., 2016; Miura et al., 2010; Quiñones-Mateu et al., 2000; Rangel et al., 2003; Troyer et al., 2005).

5.2 Conclusion

Changes within gp41 and concomitant variation in Env fusion capacity seemed to play an important role in PSV entry and IMC RC in general and that increased fusion capacity seemed a common feature of variants with high frequency. Therefore, although Env evolution seemed to follow different functional pathways, where some Env had enhanced cleavage and others had better incorporation into pseudovirus, it still culminated in high fusion capacity. Variant evolution was functionally constrained and occurred under different immune pressures within each participant so the strength of the relationship between viral diversity, Env function and disease progression varied between individuals. However, in spite of host-specific selective pressures, two phylogenetically distinct viruses in a single individual still tended to undergo early recombination within gp41 that favoured enhanced Env function leading to viral outgrowth and increased disease progression.

5.3 Summary of results

Table 5.1. Summary of Env phenotype associated with *in vivo* outgrowth overtime

Participant ID	Association of *in vivo* outgrowth at 12 mpi with:									
	Entry efficiency	RC	Coreceptor tropism	Cellular tropism	Fusion capacity	Expression	Gp160 cleavage	Gp120 incorporation	CD4 count	VL
CAP37	Yes	ND	No	No	No	ND	ND	ND	Yes	No
CAP84	No	ND	No	No	No	ND	ND	ND	No	No
CAP137	Yes	Yes	No	No	*Yes	Yes*	Yes*	No	Yes	No
CAP267	Yes	Yes	No	No	*Yes	No*	No*	No	Yes	No

For each participant, an association was determined between variant frequency and Env phenotype (entry efficiency, replication capacity, co-receptor tropism, fusion capacity, gp160 cleavage and gp120 incorporation) and disease progression (CD4+ T cell counts and viral load). Yes, means there was an association whereas No indicates that there was no association for all clones. Yes* and No* means that the outgrowth of only some clones were associated with Env phenotype.

Table 5.2. Association between fitness determinants and Env phenotype

Participant ID	Fitness determinant	Mutant generated	Entry efficiency	Fusion capacity	Expression	Gp120 incorporation	Replication capacity
CAP84	Gp41	Gp120/gp41chimeras (84c1/c4 and 84c4/c1)	Yes	No	ND	ND	ND
	PNG at N339	Introduction of N339 into the fit 84c1/c4	Yes	ND	ND	ND	ND
		Deletion of N339 in unfit 84c4	No	ND	ND	ND	ND
Cap137	PNG at N332	Deletion of N332 in fit 137c10	Yes	ND	No	No	No
CAP267	5 a.a. in CT of gp41 and 1 a.a. in C5 of gp120	Gp120/gp41chimeras (267c2/c6 and 267c6/c2)	Yes	Yes	No	No	No

Clones from CAP84, CAP137 and CAP267 were mutated and each mutant was tested for changes in Envelope phenotype. Yes, means that when the Env was mutated then phenotype changes and No means that mutations did not impact Env phenotype.

BIBILOGRAPHY

Abrahams, M.-R., Anderson, J.A., Giorgi, E.E., Seoighe, C., Mlisana, K., Ping, L.-H., Athreya, G.S., et al. (2009), "Quantitating the Multiplicity of Infection with Human Immunodeficiency Virus Type 1 Subtype C Reveals a Non-Poisson Distribution of Transmitted Variants", *Journal of Virology*, Vol. 83 No. 8, pp. 3556–3567.

Abrahams, M.-R., Treurnicht, F.K., Ngandu, N.K., Goodier, S.A., Marais, J.C., Bredell, H., Thebus, R., et al. (2013), "Rapid, complex adaption of transmitted HIV-1 full-length genomes in subtype C-infected individuals with differing disease progression", *AIDS (London, England)*, Vol. 27 No. 4, p. 507.

Abrahamyan, L.G., Mkrtchyan, S.R., Binley, J., Lu, M., Melikyan, G.B. and Cohen, F.S. (2005), "The Cytoplasmic Tail Slows the Folding of Human Immunodeficiency Virus Type 1 Env from a Late Prebundle Configuration into the Six-Helix Bundle", *Journal of Virology*, Vol. 79 No. 1, pp. 106–115.

Abram, M.E., Ferris, A.L., Shao, W., Alvord, W.G. and Hughes, S.H. (2010), "Nature, Position, and Frequency of Mutations Made in a Single Cycle of HIV-1 Replication", *Journal of Virology*, Vol. 84 No. 19, pp. 9864–9878.

Affranchino, J. and González, S. (2014), "Understanding the Process of Envelope Glycoprotein Incorporation into Virions in Simian and Feline Immunodeficiency Viruses", *Viruses*, Vol. 6 No. 1, pp. 264–283.

Alexander, M., Lynch, R., Mulenga, J., Allen, S., Derdeyn, C.A. and Hunter, E. (2010), "Donor and Recipient Envs from Heterosexual Human Immunodeficiency Virus Subtype C Transmission Pairs Require High Receptor Levels for Entry", *Journal of Virology*, Vol. 84 No. 8, pp. 4100–4104.

Alizon, S., von Wyl, V., Stadler, T., Kouyos, R.D., Yerly, S., Hirschel, B., Böni, J., et al. (2010), "Phylogenetic Approach Reveals That Virus Genotype Largely Determines HIV Set-Point Viral Load", edited by Wilke, C.O. *PLoS Pathogens*, Vol. 6 No. 9, p. e1001123.

Altfeld, M., Rosenberg, E.S., Shankarappa, R., Mukherjee, J.S., Hecht, F.M., Eldridge, R.L., Addo, M.M., et al. (2001), "Cellular Immune Responses and Viral Diversity in Individuals Treated during Acute and Early HIV-1 Infection", *The Journal of Experimental Medicine*, Vol. 193 No. 2, p. 169.

Anastassopoulou, C.G., Ketas, T.J., Klasse, P.J. and Moore, J.P. (2009), "Resistance to CCR5 inhibitors caused by sequence changes in the fusion peptide of HIV-1 gp41", *Proceedings of the National Academy of Sciences*, Vol. 106 No. 13, pp. 5318–5323.

Anastassopoulou, C.G., Marozsan, A.J., Matet, A., Snyder, A.D., Arts, E.J., Kuhmann, S.E. and Moore, J.P. (2007), "Escape of HIV-1 from a Small Molecule CCR5 Inhibitor Is Not Associated with a Fitness Loss", *PLoS Pathogens*, Vol. 3 No. 6, p. e79.

Andreani, G., Espada, C., Ceballos, A., Ambrosioni, J., Petroni, A., Pugliese, D., Bouzas, M.B., et al. (2011), "Detection of HIV-1 dual infections in highly exposed treated patients", *Virology Journal*, Vol. 8 No. 1, p. 1.

Arenas, M., Lorenzo-Redondo, R. and Lopez-Galindez, C. (2016), "Influence of mutation and recombination on HIV-1 in vitro fitness recovery", *Molecular Phylogenetics and Evolution*, Vol. 94, pp. 264–270.

Arnott, A., Jardine, D., Wilson, K., Gorry, P.R., Merlin, K., Grey, P., Law, M.G., et al. (2010), "High viral fitness during acute HIV-1 infection", *PloS One*, Vol. 5 No. 9, p. e12631.

Artenstein, A.W., VanCott, T.C., Mascola, J.R., Carr, J.K., Hegerich, P.A., Gaywee, J., Sanders-Buell, E., et al. (1995), "Dual Infection with Human Immunodeficiency Virus Type 1 of Distinct Envelope Subtypes in Humans", *Journal of Infectious Diseases*, Vol. 171 No. 4, pp. 805–810.

Arthos, J., Cicala, C., Martinelli, E., Macleod, K., Van Ryk, D., Wei, D., Xiao, Z., et al. (2008), "HIV-1 envelope protein binds to and signals through integrin α4β7, the gut mucosal homing receptor for peripheral T cells", *Nature Immunology*, Vol. 9 No. 3, pp. 301–309.

Asmal, M., Hellmann, I., Liu, W., Keele, B.F., Perelson, A.S., Bhattacharya, T., Gnanakaran, S., et al. (2011), "A Signature in HIV-1 Envelope Leader Peptide Associated with Transition from Acute to Chronic Infection Impacts Envelope Processing and Infectivity", edited by Vartanian, J.-P. *PLoS ONE*, Vol. 6 No. 8, p. e23673.

Baatz, F., Nijhuis, M., Lemaire, M., Riedijk, M., Wensing, A.M.J., Servais, J.-Y., van Ham, P.M., et al. (2011), "Impact of the HIV-1 env Genetic Context outside HR1–HR2 on Resistance to the Fusion Inhibitor Enfuvirtide and Viral Infectivity in Clinical Isolates", edited by Marais, B.J. *PLoS ONE*, Vol. 6 No. 7, p. e21535.

Bachrach, E., Dreja, H., Lin, Y.-L., Mettling, C., Pinet, V., Corbeau, P. and Piechaczyk, M. (2005), "Effects of virion surface gp120 density on infection by HIV-1 and viral production by infected cells", *Virology*, Vol. 332 No. 1, pp. 418–429.

Bagaya, B.S., Vega, J.F., Tian, M., Nickel, G.C., Li, Y., Krebs, K.C., Arts, E.J., et al. (2015), "Functional bottlenecks for generation of HIV-1 intersubtype Env recombinants", *Retrovirology*, Vol. 12 No. 1, p. 1.

Bagnarelli, P., Fiorelli, L., Vecchi, M., Monachetti, A., Menzo, S. and Clementi, M. (2003), "Analysis of the functional relationship between V3 loop and gp120 context with regard to human immunodeficiency virus coreceptor usage using naturally selected sequences and different viral backbones", *Virology*, Vol. 307 No. 2, pp. 328–340.

Ball, S.C., Abraha, A., Collins, K.R., Marozsan, A.J., Baird, H., Quiñones-Mateu, M.E., Penn-Nicholson, A., et al. (2003), "Comparing the ex vivo fitness of CCR5-tropic human immunodeficiency virus type 1 isolates of subtypes B and C.", *Journal of Virology*, Vol. 77 No. 2, pp. 1021–38.

Baribaud, F., Pöhlmann, S. and Doms, R.W. (2001), "The Role of DC-SIGN and DC-SIGNR in HIV and SIV Attachment, Infection, and Transmission", *Virology*, Vol. 286 No. 1, pp. 1–6.

Barouch, D.H. and Letvin, N.L. (2002), "Viral evolution and challenges in the development of HIV vaccines", *Vaccine*, Vol. 20, pp. A66–A68.

Barouch, D.H., O'Brien, K.L., Simmons, N.L., King, S.L., Abbink, P., Maxfield, L.F., Sun, Y.-H., et al. (2010), "Mosaic HIV-1 vaccines expand the breadth and depth of cellular immune responses in rhesus monkeys", *Nature Medicine*, Vol. 16 No. 3, pp. 319–323.

Beauparlant, D., Rusert, P., Magnus, C., Kadelka, C., Weber, J., Uhr, T., Zagordi, O., et al. (2017), "Delineating CD4 dependency of HIV-1: Adaptation to infect low level CD4 expressing target cells widens cellular tropism but severely impacts on envelope functionality", edited by Evans, D.T. *PLOS Pathogens*, Vol. 13 No. 3, p. e1006255.

Bertels, F., Marzel, A., Leventhal, G., Mitov, V., Fellay, J., Günthard, H.F., Böni, J., et al. (2017), "Dissecting HIV Virulence: Heritability of Setpoint Viral Load, CD4+ T Cell Decline and Per-Parasite Pathogenicity", *Molecular Biology and Evolution*, Vol. 35 No. 1, pp. 27–37.

Bhakta, S.J., Shang, L., Prince, J.L., Claiborne, D.T. and Hunter, E. (2011), "Mutagenesis of tyrosine and di-leucine motifs in the HIV-1 envelope cytoplasmic domain results in a loss of Env-mediated fusion and infectivity", *Retrovirology*, Vol. 8 No. 1, p. 1.

Binley, J.M., Ban, Y.E.A., Crooks, E.T., Eggink, D., Osawa, K., Schief, W.R. and Sanders, R.W. (2010), "Role of Complex Carbohydrates in Human Immunodeficiency Virus Type 1 Infection and Resistance to Antibody Neutralization", *Journal of Virology*, Vol. 84 No. 11, pp. 5637–5655.

Blaak, H., Ran, L.J., Rientsma, R. and Schuitemaker, H. (2000), "Susceptibility of in Vitro Stimulated PBMC to Infection with NSI HIV-1 Is Associated with Levels of CCR5 Expression and β-Chemokine Production", *Virology*, Vol. 267 No. 2, pp. 237–246.

Blackard, J.T., Cohen, D.E. and Mayer, K.H. (2002), "Human immunodeficiency virus superinfection and recombination: current state of knowledge and potential clinical consequences", *Clinical Infectious Diseases*, Vol. 34 No. 8, pp. 1108–1114.

Blay, W.M., Kasprzyk, T., Misher, L., Richardson, B.A. and Haigwood, N.L. (2007), "Mutations in Envelope gp120 Can Impact Proteolytic Processing of the gp160 Precursor and Thereby Affect Neutralization Sensitivity of Human Immunodeficiency Virus Type 1 Pseudoviruses", *Journal of Virology*, Vol. 81 No. 23, pp. 13037–13049.

Brenchley, J.M., Schacker, T.W., Ruff, L.E., Price, D.A., Taylor, J.H., Beilman, G.J., Nguyen, P.L., et al. (2004), "CD4 $^{+}$ T Cell Depletion during all Stages of HIV Disease Occurs Predominantly in the Gastrointestinal Tract", *The Journal of Experimental Medicine*, Vol. 200 No. 6, pp. 749–759.

Brenner, B., Routy, J.-P., Quan, Y., Moisi, D., Oliveira, M., Turner, D. and Wainberg, M.A. (2004), "Persistence of multidrug-resistant HIV-1 in primary infection leading to superinfection", *AIDS*, Vol. 18 No. 12, pp. 1653–1660.

Brockman, M.A., Schneidewind, A., Lahaie, M., Schmidt, A., Miura, T., DeSouza, I., Ryvkin, F., et al. (2007), "Escape and compensation from early HLA-B57-mediated cytotoxic T-lymphocyte pressure on human immunodeficiency virus type 1 Gag alter capsid interactions with cyclophilin A", *Journal of Virology*, Vol. 81 No. 22, pp. 12608–12618.

Bunnik, E.M., Lobbrecht, M.S.D., van Nuenen, A.C. and Schuitemaker, H. (2010), "Escape from autologous humoral immunity of HIV-1 is not associated with a decrease in replicative capacity", *Virology*, Vol. 397 No. 1, pp. 224–230.

Burke, D.S. (1997), "Recombination in HIV: an important viral evolutionary strategy.", *Emerging Infectious Diseases*, Vol. 3 No. 3, p. 253.

Burton, D.R. and Mascola, J.R. (2015), "Antibody responses to envelope glycoproteins in HIV-1 infection", *Nature Immunology*, Vol. 16 No. 6, pp. 571–576.

Bushman, F.D., Hoffmann, C., Ronen, K., Malani, N., Minkah, N., Rose, H.M., Tebas, P., et al. (2008), "Massively parallel pyrosequencing in HIV research", *AIDS*, Vol. 22 No. 12, pp. 1411–1415.

Campbell, T.B., Schneider, K., Wrin, T., Petropoulos, C.J. and Connick, E. (2003), "Relationship between In Vitro Human Immunodeficiency Virus Type 1 Replication Rate and Virus Load in Plasma", *Journal of Virology*, Vol. 77 No. 22, pp. 12105–12112.

Capel, E., Martrus, G., Parera, M., Clotet, B. and Martinez, M.A. (2012), "Evolution of the human immunodeficiency virus type 1 protease: effects on viral replication capacity and protease robustness", *Journal of General Virology*, Vol. 93 No. Pt_12, pp. 2625–2634.

Casado, C., Pernas, M., Alvaro, T., Sandonis, V., García, S., Rodríguez, C., Romero, J. del, et al. (2007), "Coinfection and Superinfection in Patients with Long-Term, Nonprogressive HIV-1 Disease", *The Journal of Infectious Diseases*, Vol. 196 No. 6, pp. 895–899.

Castro, E., Recordon-Pinson, P., Cavassini, M. and Fleury, H. (2012), "Multiclass primary antiretroviral drug resistance in a patient presenting HIV-1/2 dual infection", *Antiviral Therapy*, Vol. 17 No. 3, pp. 593–594.

Cavrois, M., Neidleman, J., Santiago, M.L., Derdeyn, C.A., Hunter, E. and Greene, W.C. (2014), "Enhanced Fusion and Virion Incorporation for HIV-1 Subtype C Envelope Glycoproteins with Compact V1/V2 Domains", *Journal of Virology*, Vol. 88 No. 4, pp. 2083–2094.

Chan, D.C., Fass, D., Berger, J.M. and Kim, P.S. (1997), "Core structure of gp41 from the HIV envelope glycoprotein", *Cell*, Vol. 89 No. 2, pp. 263–273.

Charpentier, C., Nora, T., Tenaillon, O., Clavel, F. and Hance, A.J. (2006), "Extensive Recombination among Human Immunodeficiency Virus Type 1 Quasispecies Makes an Important Contribution to Viral Diversity in Individual Patients", *Journal of Virology*, Vol. 80 No. 5, pp. 2472–2482.

Chatziandreou, N., Arauz, A.B., Freitas, I., Nyein, P.H., Fenton, G., Mehta, S.H., Kirk, G.D., et al. (2012), "Sensitivity Changes over the Course of Infection Increases the Likelihood of Resistance Against Fusion but Not CCR5 Receptor Blockers", *AIDS Research and Human Retroviruses*, Vol. 28 No. 12, pp. 1584–1593.

Checkley, M.A., Luttge, B.G. and Freed, E.O. (2011), "HIV-1 Envelope Glycoprotein Biosynthesis, Trafficking, and Incorporation", *Journal of Molecular Biology*, Vol. 410 No. 4, pp. 582–608.

Chertova, E., Bess, J.W., Crise, B.J., Sowder II, R.C., Schaden, T.M., Hilburn, J.M., Hoxie, J.A., et al. (2002), "Envelope Glycoprotein Incorporation, Not Shedding of Surface Envelope Glycoprotein (gp120/SU), Is the Primary Determinant of SU Content of Purified Human Immunodeficiency Virus Type 1 and Simian Immunodeficiency Virus", *Journal of Virology*, Vol. 76 No. 11, pp. 5315–5325.

Chikere, K., Chou, T., Gorry, P.R. and Lee, B. (2013), "Affinofile profiling: how efficiency of CD4/CCR5 usage impacts the biological and pathogenic phenotype of HIV.", *Virology*, Vol. 435 No. 1, pp. 81–91.

Chikere, K., Webb, N.E., Chou, T., Borm, K., Sterjovski, J., Gorry, P.R. and Lee, B. (2014), "Distinct HIV-1 entry phenotypes are associated with transmission, subtype specificity, and resistance to broadly neutralizing antibodies", *Retrovirology*, Vol. 11 No. 1, p. 48.

Chohan, B., Lang, D., Sagar, M., Korber, B., Lavreys, L., Richardson, B. and Overbaugh, J. (2005), "Selection for Human Immunodeficiency Virus Type 1 Envelope Glycosylation Variants with Shorter V1-V2 Loop Sequences Occurs during Transmission of Certain Genetic Subtypes and May Impact Viral RNA Levels", *Journal of Virology*, Vol. 79 No. 10, pp. 6528–6531.

Chopera, D.R., Woodman, Z., Mlisana, K., Mlotshwa, M., Martin, D.P., Seoighe, C., Treurnicht, F., et al. (2008), "Transmission of HIV-1 CTL escape variants provides HLA-mismatched recipients with a survival advantage.", *PLoS Pathogens*, Vol. 4 No. 3, p. e1000033.

Cilliers, T., Moore, P., Coetzer, M. and Morris, L. (2005), "In vitro generation of HIV type 1 subtype C isolates resistant to enfuvirtide", *AIDS Research & Human Retroviruses*, Vol. 21 No. 9, pp. 776–783.

Claiborne, D.T., Prince, J.L., Scully, E., Macharia, G., Micci, L., Lawson, B., Kopycinski, J., et al. (2015), "Replicative fitness of transmitted HIV-1 drives acute immune activation, proviral load in memory CD4+ T cells, and disease progression", *Proceedings of the National Academy of Sciences of the United States of America*, Vol. 112 No. 12, pp. E1480-1489.

Coffin, J. and Swanstrom, R. (2013), "HIV Pathogenesis: Dynamics and Genetics of Viral Populations and Infected Cells", *Cold Spring Harbor Perspectives in Medicine*, Vol. 3 No. 1, pp. a012526–a012526.

Cohen, M.S., Shaw, G.M., McMichael, A.J. and Haynes, B.F. (2011), "Acute HIV-1 Infection", *New England Journal of Medicine*, Vol. 364 No. 20, pp. 1943–1954.

Cong, M. -e., Heneine, W. and Garcia-Lerma, J.G. (2007), "The Fitness Cost of Mutations Associated with Human Immunodeficiency Virus Type 1 Drug Resistance Is Modulated by Mutational Interactions", *Journal of Virology*, Vol. 81 No. 6, pp. 3037–3041.

Connor, R.I. and Ho, D.D. (1994), "Human immunodeficiency virus type 1 variants with increased replicative capacity develop during the asymptomatic stage before disease progression.", *Journal of Virology*, Vol. 68 No. 7, pp. 4400–4408.

Connor, R.I., Sheridan, K.E., Ceradini, D., Choe, S. and Landau, N.R. (1997), "Change in coreceptor use correlates with disease progression in HIV-1–infected individuals", *Journal of Experimental Medicine*, Vol. 185 No. 4, pp. 621–628.

Cornelissen, M., Jurriaans, S., Kozaczynska, K., Prins, J.M., Hamidjaja, R.A., Zorgdrager, F., Bakker, M., et al. (2007), "Routine HIV-1 genotyping as a tool to identify dual infections", *Aids*, Vol. 21 No. 7, pp. 807–811.

Cornelissen, M., Pasternak, A.O., Grijsen, M.L., Zorgdrager, F., Bakker, M., Blom, P., Prins, J.M., et al. (2012), "HIV-1 Dual Infection Is Associated With Faster CD4+ T-Cell Decline in a Cohort of Men With Primary HIV Infection", *Clinical Infectious Diseases*, Vol. 54 No. 4, pp. 539–547.

Costa, L.J., Mayer, A.J., Busch, M.P. and Diaz, R.S. (2004), "Evidence for Selection of more Adapted Human Immunodeficiency Virus Type 1 Recombinant Strains in a Dually Infected Transfusion Recipient", *Virus Genes*, Vol. 28 No. 3, pp. 259–272.

Crawford, H., Lumm, W., Leslie, A., Schaefer, M., Boeras, D., Prado, J.G., Tang, J., et al. (2009), "Evolution of HLA-B*5703 HIV-1 escape mutations in HLA-B*5703–positive individuals and their transmission recipients", *The Journal of Experimental Medicine*, Vol. 206 No. 4, pp. 909–921.

Crawford, H., Prado, J.G., Leslie, A., Hue, S., Honeyborne, I., Reddy, S., van der Stok, M., et al. (2007), "Compensatory Mutation Partially Restores Fitness and Delays Reversion of Escape Mutation within the Immunodominant HLA-B*5703-Restricted Gag Epitope in Chronic Human Immunodeficiency Virus Type 1 Infection", *Journal of Virology*, Vol. 81 No. 15, pp. 8346–8351.

Da Silva, E.S., Mulinge, M. and Bercoff, D.P. (2013), "The frantic play of the concealed HIV envelope cytoplasmic tail", *Retrovirology*, Vol. 10 No. 1, p. 1.

da Silva, E.S., Mulinge, M., Lemaire, M., Masquelier, C., Beraud, C., Rybicki, A., Servais, J.-Y., et al. (2016), "The envelope cytoplasmic tail of HIV-1 subtype C contributes to poor replication capacity through low viral infectivity and cell-to-cell transmission", *PloS One*, Vol. 11 No. 9, p. e0161596.

Davis, M.R., Jiang, J., Zhou, J., Freed, E.O. and Aiken, C. (2006), "A Mutation in the Human Immunodeficiency Virus Type 1 Gag Protein Destabilizes the Interaction of the Envelope Protein Subunits gp120 and gp41", *Journal of Virology*, Vol. 80 No. 5, pp. 2405–2417.

De Oliveira, F., Mourez, T., Vessiere, A., Ngoupo, P.-A., Alessandri-Gradt, E., Simon, F., Rousset, D., et al. (2017), "Multiple HIV-1/M + HIV-1/O dual infections and new HIV-1/MO inter-group recombinant forms detected in Cameroon", *Retrovirology*, Vol. 14, p. 1.

Delwart, E.L., Sheppard, H.W., Walker, B.D., Goudsmit, J. and Mullins, J.I. (1994), "Human immunodeficiency virus type 1 evolution in vivo tracked by DNA heteroduplex mobility assays.", *Journal of Virology*, Vol. 68 No. 10, pp. 6672–6683.

Derdeyn, C.A., Decker, J.M., Bibollet-Ruche, F., Mokili, J.L., Muldoon, M., Denham, S.A., Heil, M.L., et al. (2004), "Envelope-Constrained Neutralization-Sensitive HIV-1 After Heterosexual Transmission", *Science*, Vol. 303 No. 5666, pp. 2019–2022.

Desrosiers, R.C. (1999), "Strategies used by human immunodeficiency virus that allow persistent viral replication", *Nature Medicine*, Vol. 5 No. 7, pp. 723–725.

Diaz-Aguilar, B., DeWispelaere, K., Yi, H.A. and Jacobs, A. (2013), "Significant Differences in Cell-Cell Fusion and Viral Entry Between Strains Revealed By Scanning Mutagenesis of The C-Heptad Repeat of HIV GP41", *Biochemistry*, Vol. 52 No. 20, pp. 3552–3563.

Didigu, C.A. and Doms, R.W. (2012), "Novel Approaches to Inhibit HIV Entry", *Viruses*, Vol. 4 No. 12, pp. 309–324.

Doitsh, G. and Greene, W.C. (2016), "Dissecting How CD4 T Cells Are Lost During HIV Infection", *Cell Host & Microbe*, Vol. 19 No. 3, pp. 280–291.

Domingo, E. and Holland, J.J. (1997), "RNA virus mutations and fitness for survival", *Annual Reviews in Microbiology*, Vol. 51 No. 1, pp. 151–178.

Doores, K.J. (2015), "The HIV glycan shield as a target for broadly neutralizing antibodies", *FEBS Journal*, Vol. 282 No. 24, pp. 4679–4691.

Dudley, D., Gao*, Y., Nelson, K., Henry, K., Nankya, I., Gibson, R. and Arts, E. (2009), "A novel yeast-based recombination method to clone and propagate diverse HIV-1 isolates", *BioTechniques*, Vol. 46 No. 6, pp. 458–467.

Dykes, C. and Demeter, L.M. (2007), "Clinical significance of human immunodeficiency virus type 1 replication fitness", *Clinical Microbiology Reviews*, Vol. 20 No. 4, pp. 550–578.

Eriksson, N., Pachter, L., Mitsuya, Y., Rhee, S.-Y., Wang, C., Gharizadeh, B., Ronaghi, M., et al. (2008), "Viral Population Estimation Using Pyrosequencing", edited by Tesler, G. *PLoS Computational Biology*, Vol. 4 No. 5, p. e1000074.

Etemad, B., Fellows, A., Kwambana, B., Kamat, A., Feng, Y., Lee, S. and Sagar, M. (2009), "Human Immunodeficiency Virus Type 1 V1-to-V5 Envelope Variants from the

Chronic Phase of Infection Use CCR5 and Fuse More Efficiently than Those from Early after Infection", *Journal of Virology*, Vol. 83 No. 19, pp. 9694–9708.

Etemad, B., Ghulam-Smith, M., Gonzalez, O., White, L.F. and Sagar, M. (2015), "Single genome amplification and standard bulk PCR yield HIV-1 envelope products with similar genotypic and phenotypic characteristics", *Journal of Virological Methods*, Vol. 214, pp. 46–53.

Fang, G., Weiser, B., Kuiken, C., Philpott, S.M., Rowland-Jones, S., Plummer, F., Kimani, J., et al. (2004), "Recombination following superinfection by HIV-1", *Aids*, Vol. 18 No. 2, pp. 153–159.

Fauci, A.S. (2007), "Pathogenesis of HIV Disease: Opportunities for New Prevention Interventions", *Clinical Infectious Diseases*, Vol. 45 No. Supplement 4, pp. S206–S212.

Frahm, N., Kiepiela, P., Adams, S., Linde, C.H., Hewitt, H.S., Sango, K., Feeney, M.E., et al. (2006), "Control of human immunodeficiency virus replication by cytotoxic T lymphocytes targeting subdominant epitopes", *Nature Immunology*, Vol. 7 No. 2, pp. 173–178.

Frankel, A.D. and Young, J.A.T. (1998), "HIV-1: Fifteen Proteins and an RNA", *Annual Review of Biochemistry*, Vol. 67 No. 1, pp. 1–25.

Freed, E.O. and Martin, M.A. (1995a), "The role of human immunodeficiency virus type 1 envelope glycoproteins in virus infection", *Journal of Biological Chemistry*, Vol. 270 No. 41, pp. 23883–23886.

Freed, E.O. and Martin, M.A. (1995b), "Virion incorporation of envelope glycoproteins with long but not short cytoplasmic tails is blocked by specific, single amino acid substitutions in the human immunodeficiency virus type 1 matrix.", *Journal of Virology*, Vol. 69 No. 3, pp. 1984–1989.

Freed, E.O. and Martin, M.A. (1996), "Domains of the human immunodeficiency virus type 1 matrix and gp41 cytoplasmic tail required for envelope incorporation into virions.", *Journal of Virology*, Vol. 70 No. 1, pp. 341–351.

Freed, E.O., Myers, D.J. and Risser, R. (1989), "Mutational analysis of the cleavage sequence of the human immunodeficiency virus type 1 envelope glycoprotein precursor gp160.", *Journal of Virology*, Vol. 63 No. 11, pp. 4670–4675.

Freed, E.O., Myers, D.J. and Risser, R. (1991), "Identification of the principal neutralizing determinant of human immunodeficiency virus type 1 as a fusion domain.", *Journal of Virology*, Vol. 65 No. 1, pp. 190–194.

Frost, S.D., Wrin, T., Smith, D.M., Pond, S.L.K., Liu, Y., Paxinos, E., Chappey, C., et al. (2005), "Neutralizing antibody responses drive the evolution of human immunodeficiency virus type 1 envelope during recent HIV infection", *Proceedings of the National Academy of Sciences of the United States of America*, Vol. 102 No. 51, pp. 18514–18519.

Fultz, P.N., Srinivasan, A., Greene, C.R., Butler, D., Swenson, R.B. and McClure, H.M. (1987), "Superinfection of a chimpanzee with a second strain of human immunodeficiency virus.", *Journal of Virology*, Vol. 61 No. 12, pp. 4026–4029.

Gaines, H., von Sydow, M., Parry, J.V., Forsgren, M., Pehrson, P.O., Sönnerborg, A., Mortimer, P.P., et al. (1988), "Detection of immunoglobulin M antibody in primary human immunodeficiency virus infection", *AIDS (London, England)*, Vol. 2 No. 1, pp. 11–15.

Garg, H. and Joshi, A. (2017), "Host and Viral Factors in HIV-Mediated Bystander Apoptosis", *Viruses*, Vol. 9 No. 8, p. 237.

Garg, H., Lee, R.T.C., Maurer-Stroh, S. and Joshi, A. (2016), "HIV-1 adaptation to low levels of CCR5 results in V3 and V2 loop changes that increase envelope pathogenicity, CCR5 affinity and decrease susceptibility to Maraviroc", *Virology*, Vol. 493, pp. 86–99.

Geijtenbeek, T.B.H., van Duijnhoven, G.C.F., van Vliet, S.J., Krieger, E., Vriend, G., Figdor, C.G. and van Kooyk, Y. (2002), "Identification of Different Binding Sites in the Dendritic Cell-specific Receptor DC-SIGN for Intercellular Adhesion Molecule 3 and HIV-1", *Journal of Biological Chemistry*, Vol. 277 No. 13, pp. 11314–11320.

Georgoulias, V. (1988), "HIV-1 and HIV-2 Double Infection In Greece", *Annals of Internal Medicine*, Vol. 108 No. 1, p. 155.

Girard, M., Osmanov, S. and Kieny, M. (2006), "A review of vaccine research and development: The human immunodeficiency virus (HIV)?", *Vaccine*, Vol. 24 No. 19, pp. 4062–4081.

Goepfert, P.A., Lumm, W., Farmer, P., Matthews, P., Prendergast, A., Carlson, J.M., Derdeyn, C.A., et al. (2008), "Transmission of HIV-1 Gag immune escape mutations is associated with reduced viral load in linked recipients", *The Journal of Experimental Medicine*, Vol. 205 No. 5, pp. 1009–1017.

Gonzales, M.J., Delwart, E., Rhee, S.-Y., Tsui, R., Zolopa, A.R., Taylor, J. and Shafer, R.W. (2003), "Lack of detectable human immunodeficiency virus type 1 superinfection during 1072 person-years of observation", *Journal of Infectious Diseases*, Vol. 188 No. 3, pp. 397–405.

Goonetilleke, N., Liu, M.K., Salazar-Gonzalez, J.F., Ferrari, G., Giorgi, E., Ganusov, V.V., Keele, B.F., et al. (2009), "The first T cell response to transmitted/founder virus contributes to the control of acute viremia in HIV-1 infection", *Journal of Experimental Medicine*, Vol. 206 No. 6, pp. 1253–1272.

Gordon, K., Omar, S., Nofemela, A., Bandawe, G., Williamson, C. and Woodman, Z. (2016), "Short Communication: A Recombinant Variant with Increased Envelope Entry Efficiency Emerged During Early Infection of an HIV-1 Subtype C Dual Infected Rapid Progressor", *AIDS Research and Human Retroviruses*, Vol. 32 No. 3, pp. 303–310.

Gorry, P.R. and Ancuta, P. (2011), "Coreceptors and HIV-1 Pathogenesis", *Current HIV/AIDS Reports*, Vol. 8 No. 1, pp. 45–53.

Gorry, P.R., Francella, N., Lewin, S.R. and Collman, R.G. (2014), "HIV-1 envelope–receptor interactions required for macrophage infection and implications for current HIV-1 cure strategies", *Journal of Leukocyte Biology*, Vol. 95 No. 1, pp. 71–81.

Gorry, P.R., Sterjovski, J., Churchill, M., Witlox, K., Gray, L., Cunningham, A. and Wesselingh, S. (2004), "The role of viral coreceptors and enhanced macrophage tropism in human immunodeficiency virus type 1 disease progression", *Sexual Health*, Vol. 1 No. 1, pp. 23–34.

Gorry, P.R., Taylor, J., Holm, G.H., Mehle, A., Morgan, T., Cayabyab, M., Farzan, M., et al. (2002), "Increased CCR5 affinity and reduced CCR5/CD4 dependence of a neurovirulent primary human immunodeficiency virus type 1 isolate", *Journal of Virology*, Vol. 76 No. 12, pp. 6277–6292.

Gottlieb, G.S., Nickle, D.C., Jensen, M.A., Wong, K.G., Grobler, J., Li, F., Liu, S.-L., et al. (2004), "Dual HIV-1 infection associated with rapid disease progression", *Lancet (London, England)*, Vol. 363 No. 9409, pp. 619–622.

Göttlinger, H.G., Sodroski, J.G. and Haseltine, W.A. (1989), "Role of capsid precursor processing and myristoylation in morphogenesis and infectivity of human immunodeficiency virus type 1.", *Proceedings of the National Academy of Sciences of the United States of America*, Vol. 86 No. 15, pp. 5781–5785.

Graham, F.L., Smiley, J., Russell, W.C. and Nairn, R. (1977), "Characteristics of a human cell line transformed by DNA from human adenovirus type 5", *Journal of General Virology*, Vol. 36 No. 1, pp. 59–72.

Gray, L., Sterjovski, J., Churchill, M., Ellery, P., Nasr, N., Lewin, S.R., Crowe, S.M., et al. (2005), "Uncoupling coreceptor usage of human immunodeficiency virus type 1 (HIV-1) from macrophage tropism reveals biological properties of CCR5-restricted HIV-1 isolates from patients with acquired immunodeficiency syndrome", *Virology*, Vol. 337 No. 2, pp. 384–398.

Grobler, J., Gray, C.M., Rademeyer, C., Seoighe, C., Ramjee, G., Karim, S.A., Morris, L., et al. (2004), "Incidence of HIV-1 dual infection and its association with increased viral load set point in a cohort of HIV-1 subtype C-infected female sex workers", *Journal of Infectious Diseases*, Vol. 190 No. 7, pp. 1355–1359.

Haaland, R.E., Hawkins, P.A., Salazar-Gonzalez, J., Johnson, A., Tichacek, A., Karita, E., Manigart, O., et al. (2009), "Inflammatory Genital Infections Mitigate a Severe Genetic Bottleneck in Heterosexual Transmission of Subtype A and C HIV-1", edited by Trkola, A.*PLoS Pathogens*, Vol. 5 No. 1, p. e1000274.

Haase, A.T. (2010), "Targeting early infection to prevent HIV-1 mucosal transmission", *Nature*, Vol. 464 No. 7286, pp. 217–223.

Haase, A.T. (2011), "Early Events in Sexual Transmission of HIV and SIV and Opportunities for Interventions", *Annual Review of Medicine*, Vol. 62 No. 1, pp. 127–139.

Haddox, H.K., Dingens, A.S. and Bloom, J.D. (2016), "Experimental estimation of the effects of all amino-acid mutations to HIV's envelope protein on viral replication in cell culture", *PLoS Pathogens,* Vol. 12 No. 12, p. e1006114.

Hartley, O., Klasse, P.J., Sattentau, Q.J. and Moore, J.P. (2005), "V3: HIV's Switch-Hitter", *AIDS Research and Human Retroviruses,* Vol. 21 No. 2, pp. 171–189.

Hemelaar, J., Gouws, E., Ghys, P.D. and Osmanov, S. (2011), "Global trends in molecular epidemiology of HIV-1 during 2000–2007", *AIDS,* Vol. 25 No. 5, pp. 679–689.

Herbinger, K.-H., Gerhardt, M., Piyasirisilp, S., Mloka, D., Arroyo, M.A., Hoffmann, O., Maboko, L., et al. (2006), "Frequency of HIV type 1 dual infection and HIV diversity: analysis of low-and high-risk populations in Mbeya Region, Tanzania", *AIDS Research & Human Retroviruses,* Vol. 22 No. 7, pp. 599–606.

Herrera, C., Klasse, P.J., Michael, E., Kake, S., Barnes, K., Kibler, C.W., Campbell-Gardener, L., et al. (2005), "The impact of envelope glycoprotein cleavage on the antigenicity, infectivity, and neutralization sensitivity of Env-pseudotyped human immunodeficiency virus type 1 particles", *Virology,* Vol. 338 No. 1, pp. 154–172.

Hinkle, D.E., Wiersma, W. and Jurs, S.G. (2003), *Applied Statistics for the Behavioral Sciences,* 5th ed., Houghton Mifflin, Boston.

Ho, D.D., Neumann, A.U., Perelson, A.S., Chen, W., Leonard, J.M. and Markowitz, M. (1995), "Rapid turnover of plasma virions and CD4 lymphocytes in HIV-1 infection", *Nature,* Vol. 373 No. 6510, pp. 123–126.

Hoffman, T.L. and Doms, R.W. (1999), "HIV-1 envelope determinants for cell tropism and chemokine receptor use", *Molecular Membrane Biology,* Vol. 16 No. 1, pp. 57–65.

Hollingsworth, T.D., Laeyendecker, O., Shirreff, G., Donnelly, C.A., Serwadda, D., Wawer, M.J., Kiwanuka, N., et al. (2010), "HIV-1 Transmitting Couples Have Similar Viral Load Set-Points in Rakai, Uganda", edited by Holmes, E.C.*PLoS Pathogens,* Vol. 6 No. 5, p. e1000876.

Hraber, P., Korber, B.T., Lapedes, A.S., Bailer, R.T., Seaman, M.S., Gao, H., Greene, K.M., et al. (2014), "Impact of Clade, Geography, and Age of the Epidemic on HIV-1 Neutralization by Antibodies", *Journal of Virology,* Vol. 88 No. 21, pp. 12623–12643.

Hsu, M., Harouse, J.M., Gettie, A., Buckner, C., Blanchard, J. and Cheng-Mayer, C. (2003), "Increased Mucosal Transmission but Not Enhanced Pathogenicity of the CCR5-Tropic, Simian AIDS-Inducing Simian/Human Immunodeficiency Virus SHIVSF162P3 Maps to Envelope gp120", *Journal of Virology,* Vol. 77 No. 2, pp. 989–998.

Huang, C., Lam, S.N., Acharya, P., Tang, M., Xiang, S.-H., Hussan, S.S., Stanfield, R.L., et al. (2007), "Structures of the CCR5 N terminus and of a tyrosine-sulfated antibody with HIV-1 gp120 and CD4", *Science,* Vol. 317 No. 5846, pp. 1930–1934.

Huang, X., Jin, W., Hu, K., Luo, S., Du, T., Griffin, G.E., Shattock, R.J., et al. (2012), "Highly conserved HIV-1 gp120 glycans proximal to CD4-binding region affect viral

infectivity and neutralizing antibody induction", *Virology*, Vol. 423 No. 1, pp. 97–106.

Hwang, S.S., Boyle, T.J., Lyerly, H.K. and Cullen, B.R. (1991), "Identification of the envelope V3 loop as the primary determinant of cell tropism in HIV-1", *Science*, Vol. 253 No. 5015, pp. 71–74.

Janes, H., Herbeck, J.T., Tovanabutra, S., Thomas, R., Frahm, N., Duerr, A., Hural, J., et al. (2015), "HIV-1 infections with multiple founders are associated with higher viral loads than infections with single founders", *Nature Medicine*, Vol. 21 No. 10, pp. 1139–1141.

Janini, L.M., Pieniazek, D., Peralta, J.M., Schechter, M., Tanuri, A., Vicente, A.C., Torre, N.D., et al. (1996), "Identification of single and dual infections with distinct subtypes of human immunodeficiency virus type 1 by using restriction fragment length polymorphism analysis", *Virus Genes*, Vol. 13 No. 1, pp. 69–81.

Jetzt, A.E., Yu, H., Klarmann, G.J., Ron, Y., Preston, B.D. and Dougherty, J.P. (2000), "High rate of recombination throughout the human immunodeficiency virus type 1 genome", *Journal of Virology*, Vol. 74 No. 3, pp. 1234–1240.

Jiang, J. and Aiken, C. (2007), "Maturation-Dependent Human Immunodeficiency Virus Type 1 Particle Fusion Requires a Carboxyl-Terminal Region of the gp41 Cytoplasmic Tail", *Journal of Virology*, Vol. 81 No. 18, pp. 9999–10008.

Jobes, D.V., Daoust, M., Nguyen, V.T., Padua, A., Sinangil, F., Pérez-Losada, M., Crandall, K.A., et al. (2006), "Longitudinal population analysis of dual infection with recombination in two strains of HIV type 1 subtype B in an individual from a Phase 3 HIV vaccine efficacy trial", *AIDS Research & Human Retroviruses*, Vol. 22 No. 10, pp. 968–978.

Johnson, W.E. and Desrosiers, R.C. (2002), "Viral Persistence: HIV's Strategies of Immune System Evasion", *Annual Review of Medicine*, Vol. 53 No. 1, pp. 499–518.

Johnston, S.H., Lobritz, M.A., Nguyen, S., Lassen, K., Delair, S., Posta, F., Bryson, Y.J., et al. (2009), "A Quantitative Affinity-Profiling System That Reveals Distinct CD4/CCR5 Usage Patterns among Human Immunodeficiency Virus Type 1 and Simian Immunodeficiency Virus Strains", *Journal of Virology*, Vol. 83 No. 21, pp. 11016–11026.

Joseph, S.B., Arrildt, K.T., Swanstrom, A.E., Schnell, G., Lee, B., Hoxie, J.A. and Swanstrom, R. (2014), "Quantification of Entry Phenotypes of Macrophage-Tropic HIV-1 across a Wide Range of CD4 Densities", *Journal of Virology*, Vol. 88 No. 4, pp. 1858–1869.

Joseph, S.B., Swanstrom, R., Kashuba, A.D.M. and Cohen, M.S. (2015), "Bottlenecks in HIV-1 transmission: insights from the study of founder viruses", *Nature Reviews Microbiology*, Vol. 13 No. 7, pp. 414–425.

Joshi, A., Sedano, M., Beauchamp, B., Punke, E., Mulla, Z., Meza, A., Alozie, O., et al. (2016), "HIV-1 Env glycoprotein phenotype along with immune activation

determines CD4 T cell loss in HIV patients", *Journal of Immunology (Baltimore, Md.: 1950)*, Vol. 196 No. 4, pp. 1768–1779.

Jost, S., Bernard, M.-C., Kaiser, L., Yerly, S., Hirschel, B., Samri, A., Autran, B., et al. (2002), "A patient with HIV-1 superinfection", *New England Journal of Medicine*, Vol. 347 No. 10, pp. 731–736.

Jurriaans, S., Van Gemen, B., Weverling, G.J., Van Strup, D., Nara, P., Coutinho, R., Koot, M., et al. (1994), "The Natural History of HIV-1 Infection: Virus Load and Virus Phenotype Independent Determinants of Clinical Course?", *Virology*, Vol. 204 No. 1, pp. 223–233.

Kalia, V., Sarkar, S., Gupta, P. and Montelaro, R.C. (2003), "Rational Site-Directed Mutations of the LLP-1 and LLP-2 Lentivirus Lytic Peptide Domains in the Intracytoplasmic Tail of Human Immunodeficiency Virus Type 1 gp41 Indicate Common Functions in Cell-Cell Fusion but Distinct Roles in Virion Envelope Incorporation", *Journal of Virology*, Vol. 77 No. 6, pp. 3634–3646.

Kalia, V., Sarkar, S., Gupta, P. and Montelaro, R.C. (2005), "Antibody Neutralization Escape Mediated by Point Mutations in the Intracytoplasmic Tail of Human Immunodeficiency Virus Type 1 gp41", *Journal of Virology*, Vol. 79 No. 4, pp. 2097–2107.

Karnik, A., Karnik, R. and Grefen, C. (2013), "SDM-Assist software to design site-directed mutagenesis primers introducing 'silent' restriction sites.", *BMC Bioinformatics*, Vol. 14, p. 105.

Keele, B.F., Giorgi, E.E., Salazar-Gonzalez, J.F., Decker, J.M., Pham, K.T., Salazar, M.G., Sun, C., et al. (2008), "Identification and characterization of transmitted and early founder virus envelopes in primary HIV-1 infection", *Proceedings of the National Academy of Sciences*, Vol. 105 No. 21, pp. 7552–7557.

Kerina, D., Babill, S.-P. and Muller, F. (2013), "HIV Diversity and Classification, Role in Transmission", *Advances in Infectious Diseases*, Vol. 03 No. 02, pp. 146–156.

Kiepiela, P., Leslie, A.J., Honeyborne, I., Ramduth, D., Thobakgale, C., Chetty, S., Rathnavalu, P., et al. (2004), "Dominant influence of HLA-B in mediating the potential co-evolution of HIV and HLA", *Nature*, Vol. 432 No. 7018, pp. 769–775.

Kiepiela, P., Ngumbela, K., Thobakgale, C., Ramduth, D., Honeyborne, I., Moodley, E., Reddy, S., et al. (2007), "CD8+ T-cell responses to different HIV proteins have discordant associations with viral load", *Nature Medicine*, Vol. 13 No. 1, pp. 46–53.

Kiwelu, I.E., Novitsky, V., Margolin, L., Baca, J., Manongi, R., Sam, N., Shao, J., et al. (2013), "Frequent intra-subtype recombination among HIV-1 circulating in Tanzania.", *PloS One*, Vol. 8 No. 8, p. e71131.

Koelsch, K.K., Smith, D.M., Little, S.J., Ignacio, C.C., Macaranas, T.R., Brown, A.J., Petropoulos, C.J., et al. (2003), "Clade B HIV-1 superinfection with wild-type virus after primary infection with drug-resistant clade B virus.", *AIDS (London, England)*, Vol. 17 No. 7, pp. F11–6.

Kondo, E. and Göttlinger, H.G. (1996), "A conserved LXXLF sequence is the major determinant in p6gag required for the incorporation of human immunodeficiency virus type 1 Vpr.", *Journal of Virology*, Vol. 70 No. 1, pp. 159–164.

Koning, F.A., Kwa, D., Boeser-Nunnink, B., Dekker, J., Vingerhoed, J., Hiemstra, H. and Schuitemaker, H. (2003), "Decreasing Sensitivity to RANTES (Regulated on Activation, Normally T Cell–Expressed and –Secreted) Neutralization of CC Chemokine Receptor 5–Using, Non–Syncytium-Inducing Virus Variants in the Course of Human Immunodeficiency Virus Type 1 Infection", *The Journal of Infectious Diseases*, Vol. 188 No. 6, pp. 864–872.

Koning, F.A., van Rij, R.P. and Schuitemaker, H. (2002), "Biological and molecular aspects of HIV-1 coreceptor usage", *Infection*, Vol. 51, p. 64.

Koot, M. (1993), "Prognostic Value of HIV-1 Syncytium-Inducing Phenotype for Rate of CD4+ Cell Depletion and Progression to AIDS", *Annals of Internal Medicine*, Vol. 118 No. 9, p. 681.

Kraft, C.S., Basu, D., Hawkins, P.A., Hraber, P.T., Chomba, E., Mulenga, J., Kilembe, W., et al. (2012), "Timing and source of subtype-C HIV-1 superinfection in the newly infected partner of Zambian couples with disparate viruses", *Retrovirology*, Vol. 9 No. 1, p. 22.

Kraus, M.H., Parrish, N.F., Shaw, K.S., Decker, J.M., Keele, B.F., Salazar-Gonzalez, J.F., Grayson, T., et al. (2010), "A rev1-vpu polymorphism unique to HIV-1 subtype A and C strains impairs envelope glycoprotein expression from rev-vpu-env cassettes and reduces virion infectivity in pseudotyping assays.", *Virology*, Elsevier Inc., Vol. 397 No. 2, pp. 346–57.

Kuhmann, S.E. and Hartley, O. (2008), "Targeting Chemokine Receptors in HIV: A Status Report", *Annual Review of Pharmacology and Toxicology*, Vol. 48 No. 1, pp. 425–461.

Kwong, P.D., Wyatt, R., Robinson, J., Sweet, R.W., Sodroski, J. and Hendrickson, W.A. (1998), "Structure of an HIV gp120 envelope glycoprotein in complex with the CD4 receptor and a neutralizing human antibody", *Nature*, Vol. 393 No. 6686, pp. 648–659.

Lambele´, M., Labrosse, B., Roch, E., Moreau, A., Verrier, B., Barin, F., Roingeard, P., et al. (2007), "Impact of Natural Polymorphism within the gp41 Cytoplasmic Tail of Human Immunodeficiency Virus Type 1 on the Intracellular Distribution of Envelope Glycoproteins and Viral Assembly", *Journal of Virology*, Vol. 81 No. 1, pp. 125–140.

Langford, S.E., Ananworanich, J. and Cooper, D.A. (2007), "Predictors of disease progression in HIV infection: a review", *AIDS Research and Therapy*, Vol. 4 No. 1, p. 11.

Lanxon-Cookson, E.C., Swain, J.V., Manocheewa, S., Smith, R.A., Maust, B., Kim, M., Westfall, D., et al. (2013), "Factors Affecting Relative Fitness Measurements in

Pairwise Competition Assays of Human Immunodeficiency Viruses", *Journal of Virological Methods*, Vol. 194 No. 0, p. 10.1016/j.jviromet.2013.07.062.

Lassen, K.G., Lobritz, M.A., Bailey, J.R., Johnston, S., Nguyen, S., Lee, B., Chou, T., et al. (2009), "Elite Suppressor–Derived HIV-1 Envelope Glycoproteins Exhibit Reduced Entry Efficiency and Kinetics", edited by Trkola, A.*PLoS Pathogens*, Vol. 5 No. 4, p. e1000377.

Lavine, C.L., Lao, S., Montefiori, D.C., Haynes, B.F., Sodroski, J.G., Yang, X. and the NIAID Center for HIV/AIDS Vaccine Immunology (CHAVI). (2012), "High-Mannose Glycan-Dependent Epitopes Are Frequently Targeted in Broad Neutralizing Antibody Responses during Human Immunodeficiency Virus Type 1 Infection", *Journal of Virology*, Vol. 86 No. 4, pp. 2153–2164.

Leaman, D.P. and Zwick, M.B. (2013), "Increased Functional Stability and Homogeneity of Viral Envelope Spikes through Directed Evolution", edited by Trkola, A.*PLoS Pathogens*, Vol. 9 No. 2, p. e1003184.

Lebigot, S., Roingeard, P., Thibault, G., Lemiale, F., Verrier, B., Barin, F. and Brand, D. (2001), "The Transmembrane Protein of HIV-1 Primary Isolates Modulates Cell Surface Expression of Their Envelope Glycoproteins", *Virology*, Vol. 290 No. 1, pp. 136–142.

Lee, B., Sharron, M., Montaner, L.J., Weissman, D. and Doms, R.W. (1999), "Quantification of CD4, CCR5, and CXCR4 levels on lymphocyte subsets, dendritic cells, and differentially conditioned monocyte-derived macrophages", *Proceedings of the National Academy of Sciences*, Vol. 96 No. 9, pp. 5215–5220.

Lee, J.H., Andrabi, R., Su, C.-Y., Yasmeen, A., Julien, J.-P., Kong, L., Wu, N.C., et al. (2017), "A Broadly Neutralizing Antibody Targets the Dynamic HIV Envelope Trimer Apex via a Long, Rigidified, and Anionic β-Hairpin Structure", *Immunity*, Vol. 46 No. 4, pp. 690–702.

Leslie, A.J., Pfafferott, K.J., Chetty, P., Draenert, R., Addo, M.M., Feeney, M., Tang, Y., et al. (2004), "HIV evolution: CTL escape mutation and reversion after transmission", *Nature Medicine*, Vol. 10 No. 3, pp. 282–289.

Levin, J.G., Guo, J., Rouzina, Ioulia and Musier-Forsyth, K. (2005), "Nucleic Acid Chaperone Activity of HIV-1 Nucleocapsid Protein: Critical Role in Reverse Transcription and Molecular Mechanism", *Progress in Nucleic Acid Research and Molecular Biology*, Vol. 80, Academic Press, pp. 217–286.

Li, B., Gladden, A.D., Altfeld, M., Kaldor, J.M., Cooper, D.A., Kelleher, A.D. and Allen, T.M. (2007), "Rapid Reversion of Sequence Polymorphisms Dominates Early Human Immunodeficiency Virus Type 1 Evolution", *Journal of Virology*, Vol. 81 No. 1, pp. 193–201.

Li, S., Juarez, J., Alali, M., Dwyer, D., Collman, R., Cunningham, A. and Naif, H.M. (1999), "Persistent CCR5 utilization and enhanced macrophage tropism by primary blood human immunodeficiency virus type 1 isolates from advanced stages of disease and

comparison to tissue-derived isolates", *Journal of Virology*, Vol. 73 No. 12, pp. 9741–9755.

Li, Y., Luo, L., Thomas, D.Y. and Kang, C.Y. (2000), "The HIV-1 Env Protein Signal Sequence Retards Its Cleavage and Down-regulates the Glycoprotein Folding", *Virology*, Vol. 272 No. 2, pp. 417–428.

Liu, J., Bartesaghi, A., Borgnia, M.J., Sapiro, G. and Subramaniam, S. (2008), "Molecular architecture of native HIV-1 gp120 trimers", *Nature*, Vol. 455 No. 7209, pp. 109–113.

Liu, S.-L., Mittler, J.E., Nickle, D.C., Mulvania, T.M., Shriner, D., Rodrigo, A.G., Kosloff, B., et al. (2002), "Selection for Human Immunodeficiency Virus Type 1 Recombinants in a Patient with Rapid Progression to AIDS", *Journal of Virology*, Vol. 76 No. 21, pp. 10674–10684.

Liu, S.-L., Schacker, T., Musey, L., Shriner, D., McElrath, M.J., Corey, L. and Mullins, J.I. (1997), "Divergent patterns of progression to AIDS after infection from the same source: human immunodeficiency virus type 1 evolution and antiviral responses.", *Journal of Virology*, Vol. 71 No. 6, pp. 4284–4295.

Liu, Y., McNevin, J., Zhao, H., Tebit, D.M., Troyer, R.M., McSweyn, M., Ghosh, A.K., et al. (2007), "Evolution of human immunodeficiency virus type 1 cytotoxic T-lymphocyte epitopes: fitness-balanced escape.", *Journal of Virology*, Vol. 81 No. 22, pp. 12179–88.

Lobritz, M.A., Marozsan, A.J., Troyer, R.M. and Arts, E.J. (2007), "Natural Variation in the V3 Crown of Human Immunodeficiency Virus Type 1 Affects Replicative Fitness and Entry Inhibitor Sensitivity", *Journal of Virology*, Vol. 81 No. 15, pp. 8258–8269.

Lodermeyer, V., Suhr, K., Schrott, N., Kolbe, C., Stürzel, C.M., Krnavek, D., Münch, J., et al. (2013), "90K, an interferon-stimulated gene product, reduces the infectivity of HIV-1", *Retrovirology*, Vol. 10 No. 1, p. 111.

Lu, J., Sista, P., Giguel, F., Greenberg, M. and Kuritzkes, D.R. (2004), "Relative Replicative Fitness of Human Immunodeficiency Virus Type 1 Mutants Resistant to Enfuvirtide (T-20)", *Journal of Virology*, Vol. 78 No. 9, pp. 4628–4637.

Luan, H., Han, X., Yu, X., An, M., Zhang, H., Zhao, B., Xu, J., et al. (2017), "Dual Infection Contributes to Rapid Disease Progression in Men Who Have Sex With Men in China", *JAIDS Journal of Acquired Immune Deficiency Syndromes*, Vol. 75 No. 4, pp. 480–487.

Lynch, R.M., Wong, P., Tran, L., O'Dell, S., Nason, M.C., Li, Y., Wu, X., et al. (2015), "HIV-1 Fitness Cost Associated with Escape from the VRC01 Class of CD4 Binding Site Neutralizing Antibodies", edited by Doms, R.W. *Journal of Virology*, Vol. 89 No. 8, pp. 4201–4213.

Maartens, G., Celum, C. and Lewin, S.R. (2014), "HIV infection: epidemiology, pathogenesis, treatment, and prevention", *The Lancet*, Vol. 384 No. 9939, pp. 258–271.

Magiorkinis, G. (2003), “In vivo characteristics of human immunodeficiency virus type 1 intersubtype recombination: determination of hot spots and correlation with sequence similarity”, *Journal of General Virology*, Vol. 84 No. 10, pp. 2715–2722.

Mammano, F., Kondo, E., Sodroski, J., Bukovsky, A. and Göttlinger, H.G. (1995), “Rescue of human immunodeficiency virus type 1 matrix protein mutants by envelope glycoproteins with short cytoplasmic domains.”, *Journal of Virology*, Vol. 69 No. 6, pp. 3824–3830.

Mandalia, S., Westrop, S.J., Beck, E.J., Nelson, M., Gazzard, B.G. and Imami, N. (2012), “Are Long-Term Non-Progressors Very Slow Progressors? Insights from the Chelsea and Westminster HIV Cohort, 1988–2010”, edited by Landay, A. *PLoS ONE*, Vol. 7 No. 2, p. e29844.

Mani, I., Gilbert, P., Sankale, J.-L., Eisen, G., Mboup, S. and Kanki, P.J. (2002), “Intrapatient Diversity and Its Correlation with Viral Setpoint in Human Immunodeficiency Virus Type 1 CRF02_A/G-IbNG Infection”, *Journal of Virology*, Vol. 76 No. 21, pp. 10745–10755.

Mansky, L.M. and Temin, H.M. (1995), “Lower in vivo mutation rate of human immunodeficiency virus type 1 than that predicted from the fidelity of purified reverse transcriptase.”, *Journal of Virology*, Vol. 69 No. 8, pp. 5087–5094.

Mao, Y., Wang, L., Gu, C., Herschhorn, A., Xiang, S.-H., Haim, H., Yang, X., et al. (2012), “Subunit organization of the membrane-bound HIV-1 envelope glycoprotein trimer”, *Nature Structural & Molecular Biology*, Vol. 19 No. 9, pp. 893–899.

Markosyan, R.M., Cohen, F.S. and Melikyan, G.B. (2003), “HIV-1 envelope proteins complete their folding into six-helix bundles immediately after fusion pore formation”, *Molecular Biology of the Cell*, Vol. 14 No. 3, pp. 926–938.

Markosyan, R.M., Leung, M.Y. and Cohen, F.S. (2009), “The Six-Helix Bundle of Human Immunodeficiency Virus Env Controls Pore Formation and Enlargement and Is Initiated at Residues Proximal to the Hairpin Turn”, *Journal of Virology*, Vol. 83 No. 19, pp. 10048–10057.

Marozsan, A.J., Moore, D.M., Lobritz, M.A., Fraundorf, E., Abraha, A., Reeves, J.D. and Arts, E.J. (2005), “Differences in the Fitness of Two Diverse Wild-Type Human Immunodeficiency Virus Type 1 Isolates Are Related to the Efficiency of Cell Binding and Entry”, *Journal of Virology*, Vol. 79 No. 11, pp. 7121–7134.

Martin, F., Lee, J., Thomson, E., Tarrant, N., Hale, A. and Lacey, C.J. (2016), “Two cases of possible transmitted drug-resistant HIV: likely HIV superinfection and unmasking of pre-existing resistance”, *International Journal of STD & AIDS*, Vol. 27 No. 1, pp. 66–69.

Martinez-Picado, J., Prado, J.G., Fry, E.E., Pfafferott, K., Leslie, A., Chetty, S., Thobakgale, C., et al. (2006), “Fitness Cost of Escape Mutations in p24 Gag in Association with Control of Human Immunodeficiency Virus Type 1”, *Journal of Virology*, Vol. 80 No. 7, pp. 3617–3623.

McCune, J.M., Rabin, L.B., Feinberg, M.B., Lieberman, M., Kosek, J.C., Reyes, G.R. and Weissman, I.L. (1988), "Endoproteolytic cleavage of gp160 is required for the activation of human immunodeficiency virus", *Cell*, Vol. 53 No. 1, pp. 55–67.

McKEATING, J.A., McKNIGHT, A. and Moore, J.P. (1991), "Differential loss of envelope glycoprotein gp120 from virions of human immunodeficiency virus type 1 isolates: effects on infectivity and neutralization.", *Journal of Virology*, Vol. 65 No. 2, pp. 852–860.

McMichael, A.J., Borrow, P., Tomaras, G.D., Goonetilleke, N. and Haynes, B.F. (2010), "The immune response during acute HIV-1 infection: clues for vaccine development", *Nature Reviews Immunology*, Vol. 10 No. 1, pp. 11–23.

Mehandru, S., Poles, M.A., Tenner-Racz, K., Horowitz, A., Hurley, A., Hogan, C., Boden, D., et al. (2004), "Primary HIV-1 Infection Is Associated with Preferential Depletion of CD4 $^{+}$ T Lymphocytes from Effector Sites in the Gastrointestinal Tract", *The Journal of Experimental Medicine*, Vol. 200 No. 6, pp. 761–770.

Meijerink, H., Indrati, A.R., van Crevel, R., Joosten, I., Koenen, H. and van der Ven, A.J. (2014), "The number of CCR5 expressing CD4+ T lymphocytes is lower in HIV-infected long-term non-progressors with viral control compared to normal progressors: a cross-sectional study", *BMC Infectious Diseases*, Vol. 14 No. 1, p. 683.

Mellors, J.W. (1997), "Plasma Viral Load and CD4+ Lymphocytes as Prognostic Markers of HIV-1 Infection", *Annals of Internal Medicine*, Vol. 126 No. 12, p. 946.

Mellors, J.W., Rinaldo, C.R., Gupta, P., White, R.M., Todd, J.A. and Kingsley, L.A. (1996), "Prognosis in HIV-1 infection predicted by the quantity of virus in plasma", *Science (New York, N.Y.)*, Vol. 272 No. 5265, pp. 1167–1170.

Mervis, R.J., Ahmad, N., Lillehoj, E.P., Raum, M.G., Salazar, F.H., Chan, H.W. and Venkatesan, S. (1988), "The gag gene products of human immunodeficiency virus type 1: alignment within the gag open reading frame, identification of posttranslational modifications, and evidence for alternative gag precursors.", *Journal of Virology*, Vol. 62 No. 11, pp. 3993–4002.

Michael, N.L. (1999), "Host genetic influences on HIV-1 pathogenesis", *Current Opinion in Immunology*, Vol. 11 No. 4, pp. 466–474.

Milicic, A., Price, D.A., Zimbwa, P., Booth, B.L., Brown, H.L., Easterbrook, P.J., Olsen, K., et al. (2005), "CD8+ T Cell Epitope-Flanking Mutations Disrupt Proteasomal Processing of HIV-1 Nef", *The Journal of Immunology*, Vol. 175 No. 7, pp. 4618–4626.

Miura, T., Brockman, M.A., Schneidewind, A., Lobritz, M., Pereyra, F., Rathod, A., Block, B.L., et al. (2008), "HLA-B57/B* 5801 HIV-1 elite controllers select for rare gag variants associated with reduced viral replication capacity and strong Ctl recognition", *Journal of Virology*.

Miura, T., Brumme, Z.L., Brockman, M.A., Rosato, P., Sela, J., Brumme, C.J., Pereyra, F., et al. (2010), "Impaired Replication Capacity of Acute/Early Viruses in Persons Who Become HIV Controllers", *Journal of Virology*, Vol. 84 No. 15, pp. 7581–7591.

Mlisana, K., Werner, L., Garrett, N.J., McKinnon, L.R., van Loggerenberg, F., Passmore, J.-A.S., Gray, C.M., et al. (2014), "Rapid Disease Progression in HIV-1 Subtype C–Infected South African Women", *Clinical Infectious Diseases*, Vol. 59 No. 9, pp. 1322–1331.

Modrow, S., Hahn, B.H., Shaw, G.M., Gallo, R.C., Wong-Staal, F. and Wolf, H. (1987), "Computer-assisted analysis of envelope protein sequences of seven human immunodeficiency virus isolates: prediction of antigenic epitopes in conserved and variable regions.", *Journal of Virology*, Vol. 61 No. 2, pp. 570–578.

Mohri, H., Prada, N. and Markowitz, M. (2015), "The Viral Envelope is a Major Determinant of Enhanced Fitness of a Multidrug-Resistant HIV-1 Variant", *Journal of Acquired Immune Deficiency Syndromes (1999)*, Vol. 68 No. 5, p. 487.

Mondor, I., Ugolini, S. and Sattentau, Q.J. (1998), "Human immunodeficiency virus type 1 attachment to HeLa CD4 cells is CD4 independent and gp120 dependent and requires cell surface heparans", *Journal of Virology*, Vol. 72 No. 5, pp. 3623–3634.

Montero, M., van Houten, N.E., Wang, X. and Scott, J.K. (2008), "The Membrane-Proximal External Region of the Human Immunodeficiency Virus Type 1 Envelope: Dominant Site of Antibody Neutralization and Target for Vaccine Design", *Microbiology and Molecular Biology Reviews*, Vol. 72 No. 1, pp. 54–84.

Moore, C.B., John, M., James, I.R., Christiansen, F.T., Witt, C.S. and Mallal, S.A. (2002), "Evidence of HIV-1 Adaptation to HLA-Restricted Immune Responses at a Population Level", *Science*, Vol. 296 No. 5572, p. 1439.

Moore, J.P., McKeating, J.A., Weiss, R.A. and Sattentau, Q.J. (1990), "Dissociation of gp120 from HIV-1 virions induced by soluble CD4", *Science*, Vol. 250 No. 4984, pp. 1139–1142.

Moore, P.L., Crooks, E.T., Porter, L., Zhu, P., Cayanan, C.S., Grise, H., Corcoran, P., et al. (2006), "Nature of Nonfunctional Envelope Proteins on the Surface of Human Immunodeficiency Virus Type 1", *Journal of Virology*, Vol. 80 No. 5, pp. 2515–2528.

Moore, P.L., Gray, E.S., Choge, I.A., Ranchobe, N., Mlisana, K., Abdool Karim, S.S., Williamson, C., et al. (2008), "The C3-V4 Region Is a Major Target of Autologous Neutralizing Antibodies in Human Immunodeficiency Virus Type 1 Subtype C Infection", *Journal of Virology*, Vol. 82 No. 4, pp. 1860–1869.

Moore, P.L., Gray, E.S., Wibmer, C.K., Bhiman, J.N., Nonyane, M., Sheward, D.J., Hermanus, T., et al. (2012), "Evolution of an HIV glycan–dependent broadly neutralizing antibody epitope through immune escape", *Nature Medicine*, Vol. 18 No. 11, pp. 1688–1692.

Mostowy, R., Kouyos, R.D., Fouchet, D. and Bonhoeffer, S. (2011), "The role of recombination for the coevolutionary dynamics of HIV and the immune response.", *PloS One*, Vol. 6 No. 2, p. e16052.

Moulard, M. and Decroly, E. (2000), "Maturation of HIV envelope glycoprotein precursors by cellular endoproteases", *Biochimica et Biophysica Acta (BBA)-Reviews on Biomembranes*, Vol. 1469 No. 3, pp. 121–132.

Moutouh, L., Corbeil, J. and Richman, D.D. (1996), "Recombination leads to the rapid emergence of HIV-1 dually resistant mutants under selective drug pressure.", *Proceedings of the National Academy of Sciences of the United States of America*, Vol. 93 No. 12, pp. 6106–11.

Murakami, T. (2008), "Roles of the interactions between Env and Gag proteins in the HIV-1 replication cycle", *Microbiology and Immunology*, Vol. 52 No. 5, pp. 287–295.

Murakami, T. and Freed, E.O. (2000), "Genetic Evidence for an Interaction between Human Immunodeficiency Virus Type 1 Matrix and α-Helix 2 of the gp41 Cytoplasmic Tail", *Journal of Virology*, Vol. 74 No. 8, pp. 3548–3554.

Murakami, T. and Freed, E.O. (2000), "The long cytoplasmic tail of gp41 is required in a cell type-dependent manner for HIV-1 envelope glycoprotein incorporation into virions", *Proceedings of the National Academy of Sciences*, Vol. 97 No. 1, pp. 343–348.

Murin, C.D., Julien, J.-P., Sok, D., Stanfield, R.L., Khayat, R., Cupo, A., Moore, J.P., et al. (2014), "Structure of 2G12 Fab2 in Complex with Soluble and Fully Glycosylated HIV-1 Env by Negative-Stain Single-Particle Electron Microscopy", *Journal of Virology*, Vol. 88 No. 17, pp. 10177–10188.

Mylonakis, E., Paliou, M. and Rich, J.D. (2001), "Plasma Viral Load Testing in the Management of HIV Infection", *American Family Physician*, Vol. 63 No. 3, p. 483.

Nájera, R., Delgado, E., Pérez-Alvarez, L. and Thomson, M.M. (2002), "Genetic recombination and its role in the development of the HIV-1 pandemic", *Aids*, Vol. 16, pp. S3–S16.

Ndung'u, T. and Weiss, R.A. (2012), "On HIV diversity", *AIDS*, Vol. 26 No. 10, pp. 1255–1260.

Neumann, T., Hagmann, I., Lohrengel, S., Heil, M.L., Derdeyn, C.A., Kräusslich, H.-G. and Dittmar, M.T. (2005), "T20-insensitive HIV-1 from naïve patients exhibits high viral fitness in a novel dual-color competition assay on primary cells", *Virology*, Vol. 333 No. 2, pp. 251–262.

Nijhuis, M., Deeks, S. and Boucher, C. (2001), "Implications of antiretroviral resistance on viral fitness.", *Current Opinion in Infectious Diseases*, Vol. 14 No. 1, pp. 23–28.

Nofemela, A. (2013), *Characterization of Genotypic and Phenotypic Properties of Transmitted Human Immunodeficiency Virus Type 1 Variants Circulating in Mbeya Tanzania*, University of Cape Town.

Nofemela, A., Selhorst, P., Bandawe, G. and Woodman, Z. (2015), "Phenotypic characterisation of Human Immunodeficiency Virus type 1 Envelope entry efficiency of transmitted/founder variants circulating in Mbeya, Tanzania", *Medical Research Archives*, available at: http://www.journals.ke-i.org/.

Nuzzo, R. (2014), "Scientific method: statistical errors", *Nature*, Vol. 506 No. 7487, pp. 150–152.

Olivieri, K., Scoggins, R.M., Bor, Y., Matthews, A., Mark, D., Taylor, J.R., Chernauskas, D., et al. (2007), "The envelope gene is a cytopathic determinant of CCR5 tropic HIV-1", *Virology*, Vol. 358 No. 1, pp. 23–38.

Ono, A., Orenstein, J.M. and Freed, E.O. (2000), "Role of the Gag matrix domain in targeting human immunodeficiency virus type 1 assembly", *Journal of Virology*, Vol. 74 No. 6, pp. 2855–2866.

Ostrowski, M.A., Justement, S.J., Catanzaro, A., Hallahan, C.A., Ehler, L.A., Mizell, S.B., Kumar, P.N., et al. (1998), "Expression of Chemokine Receptors CXCR4 and CCR5 in HIV-1-Infected and Uninfected Individuals", *The Journal of Immunology*, Vol. 161 No. 6, pp. 3195–3201.

Pacold, M., Smith, D., Little, S., Cheng, P.M., Jordan, P., Ignacio, C., Richman, D., et al. (2010), "Comparison of Methods to Detect HIV Dual Infection", *AIDS Research and Human Retroviruses*, Vol. 26 No. 12, pp. 1291–1298.

Pancera, M., Majeed, S., Ban, Y.-E.A., Chen, L., Huang, C. -c., Kong, L., Kwon, Y.D., et al. (2010), "Structure of HIV-1 gp120 with gp41-interactive region reveals layered envelope architecture and basis of conformational mobility", *Proceedings of the National Academy of Sciences*, Vol. 107 No. 3, pp. 1166–1171.

Pancera, M., Zhou, T., Druz, A., Georgiev, I.S., Soto, C., Gorman, J., Huang, J., et al. (2014), "Structure and immune recognition of trimeric pre-fusion HIV-1 Env", *Nature*, Vol. 514 No. 7523, pp. 455–461.

Pantaleo, G. and Fauci, and A.S. (1996), "Immunopathogenesis of Hiv Infection", *Annual Review of Microbiology*, Vol. 50 No. 1, pp. 825–854.

Parrish, N.F., Gao, F., Li, H., Giorgi, E.E., Barbian, H.J., Parrish, E.H., Zajic, L., et al. (2013), "Phenotypic properties of transmitted founder HIV-1", *Proceedings of the National Academy of Sciences*, Vol. 110 No. 17, pp. 6626–6633.

Parrish, N.F., Wilen, C.B., Banks, L.B., Iyer, S.S., Pfaff, J.M., Salazar-Gonzalez, J.F., Salazar, M.G., et al. (2012), "Transmitted/Founder and Chronic Subtype C HIV-1 Use CD4 and CCR5 Receptors with Equal Efficiency and Are Not Inhibited by Blocking the Integrin α4β7", edited by Trkola, A. *PLoS Pathogens*, Vol. 8 No. 5, p. e1002686.

Pastore, C., Nedellec, R., Ramos, A., Pontow, S., Ratner, L. and Mosier, D.E. (2006), "Human Immunodeficiency Virus Type 1 Coreceptor Switching: V1/V2 Gain-of-Fitness Mutations Compensate for V3 Loss-of-Fitness Mutations", *Journal of Virology*, Vol. 80 No. 2, pp. 750–758.

Peeters, M. (2000), "Recombinant HIV sequences: their role in the global epidemic", *HIV Sequence Compendium*, Vol. 2000, pp. 1–39.

Pejchal, R., Doores, K.J., Walker, L.M., Khayat, R., Huang, P.-S., Wang, S.-K., Stanfield, R.L., et al. (2011), "A Potent and Broad Neutralizing Antibody Recognizes and Penetrates the HIV Glycan Shield", *Science*, Vol. 334 No. 6059, pp. 1097–1103.

Pernas, M., Casado, C., Arcones, C., Llano, A., Sánchez-Merino, V., Mothe, B., Vicario, J.L., et al. (2012), "Low-Replicating Viruses and Strong Anti-Viral Immune Response Associated with Prolonged Disease Control in a Superinfected HIV-1 LTNP Elite Controller", edited by Geng, E.H. *PLoS ONE*, Vol. 7 No. 2, p. e31928.

Pernas, M., Casado, C., Fuentes, R., Pérez-Elías, M.J. and López-Galíndez, C. (2006), "A dual superinfection and recombination within HIV-1 subtype B 12 years after primoinfection", *JAIDS Journal of Acquired Immune Deficiency Syndromes*, Vol. 42 No. 1, pp. 12–18.

Phillips, A. and Pezzotti, P. (2004), "Short-term risk of Aids according to current Cd4 cell count and viral load in antiretroviral drug-naive individuals and those treated in the monotherapy era.", *Aids (London, England)*, Vol. 18 No. 1, pp. 51–58.

Phillips, R.E., Rowland-Jones, S., Nixon, D.F., Gotch, F.M., Edwards, J.P., Ogunlesi, A.O., Elvin, J.G., et al. (1991), "Human immunodeficiency virus genetic variation that can escape cytotoxic T cell recognition", *Nature*, Vol. 354 No. 6353, pp. 453–459.

Ping, L.-H., Joseph, S.B., Anderson, J.A., Abrahams, M.-R., Salazar-Gonzalez, J.F., Kincer, L.P., Treurnicht, F.K., et al. (2013), "Comparison of viral Env proteins from acute and chronic infections with subtype C human immunodeficiency virus type 1 identifies differences in glycosylation and CCR5 utilization and suggests a new strategy for immunogen design", *Journal of Virology*, Vol. 87 No. 13, pp. 7218–7233.

Platt, E.J., Wehrly, K., Kuhmann, S.E., Chesebro, B. and Kabat, D. (1998), "Effects of CCR5 and CD4 cell surface concentrations on infections by macrophagetropic isolates of human immunodeficiency virus type 1", *Journal of Virology*, Vol. 72 No. 4, pp. 2855–2864.

Pollakis, G., Kang, S., Kliphuis, A., Chalaby, M.I.M., Goudsmit, J. and Paxton, W.A. (2001), "*N*- Linked Glycosylation of the HIV Type-1 gp120 Envelope Glycoprotein as a Major Determinant of CCR5 and CXCR4 Coreceptor Utilization", *Journal of Biological Chemistry*, Vol. 276 No. 16, pp. 13433–13441.

Poon, D.T., Wu, J. and Aldovini, A. (1996), "Charged amino acid residues of human immunodeficiency virus type 1 nucleocapsid p7 protein involved in RNA packaging and infectivity.", *Journal of Virology*, Vol. 70 No. 10, pp. 6607–6616.

Postler, T.S. and Desrosiers, R.C. (2013), "The Tale of the Long Tail: the Cytoplasmic Domain of HIV-1 gp41", *Journal of Virology*, Vol. 87 No. 1, pp. 2–15.

Powell, R.L.R., Kinge, T. and Nyambi, P.N. (2010), "Infection by Discordant Strains of HIV-1 Markedly Enhances the Neutralizing Antibody Response against Heterologous Virus", *Journal of Virology*, Vol. 84 No. 18, pp. 9415–9426.

Prince, J.L., Claiborne, D.T., Carlson, J.M., Schaefer, M., Yu, T., Lahki, S., Prentice, H.A., et al. (2012), "Role of transmitted Gag CTL polymorphisms in defining replicative capacity and early HIV-1 pathogenesis.", *PLoS Pathogens*, Vol. 8 No. 11, p. e1003041.

Provine, N.M., Puryear, W.B., Wu, X., Overbaugh, J. and Haigwood, N.L. (2009), "The Infectious Molecular Clone and Pseudotyped Virus Models of Human Immunodeficiency Virus Type 1 Exhibit Significant Differences in Virion Composition with Only Moderate Differences in Infectivity and Inhibition Sensitivity", *Journal of Virology*, Vol. 83 No. 17, pp. 9002–9007.

Pugach, P., Marozsan, A.J., Ketas, T.J., Landes, E.L., Moore, J.P. and Kuhmann, S.E. (2007), "HIV-1 Clones Resistant to a Small Molecule CCR5 Inhibitor Use the Inhibitor-Bound Form of CCR5 for Entry", *Virology*, Vol. 361 No. 1, pp. 212–228.

Quiñones-Mateu, M.E. and Arts, E.J. (2001), "HIV-1 fitness: implications for drug resistance, disease progression, and global epidemic evolution", *HIV Sequence Compendium*, Vol. 2001, pp. 134–170.

Quiñones-Mateu, M.E. and Arts, E.J. (2002), "Fitness of drug resistant HIV-1: methodology and clinical implications", *Drug Resistance Updates*, Vol. 5 No. 6, pp. 224–233.

Quiñones-Mateu, M.E. and Arts, E.J. (2006), "Virus Fitness: Concept, Quantification, and Application to HIV Population Dynamics", *SpringerLink*, pp. 83–140.

Quiñones-Mateu, M.E., Ball, S.C., Marozsan, A.J., Torre, V.S., Albright, J.L., Vanham, G., van der Groen, G., et al. (2000), "A dual infection/competition assay shows a correlation between ex vivo human immunodeficiency virus type 1 fitness and disease progression", *Journal of Virology*, Vol. 74 No. 19, pp. 9222–9233.

Ramirez, B.C., Simon-Loriere, E., Galetto, R. and Negroni, M. (2008), "Implications of recombination for HIV diversity", *Virus Research*, Vol. 134 No. 1–2, pp. 64–73.

Ramos, A., Hu, D.J., Nguyen, L., Phan, K.-O., Vanichseni, S., Promadej, N., Choopanya, K., et al. (2002), "Intersubtype Human Immunodeficiency Virus Type 1 Superinfection following Seroconversion to Primary Infection in Two Injection Drug Users", *Journal of Virology*, Vol. 76 No. 15, pp. 7444–7452.

Rangel, H.R., Weber, J., Chakraborty, B., Gutierrez, A., Marotta, M.L., Mirza, M., Kiser, P., et al. (2003), "Role of the Human Immunodeficiency Virus Type 1 Envelope Gene in Viral Fitness", *Journal of Virology*, Vol. 77 No. 16, pp. 9069–9073.

Ratcliff, A.N., Shi, W. and Arts, E.J. (2013), "HIV-1 Resistance to Maraviroc Conferred by a CD4 Binding Site Mutation in the Envelope Glycoprotein gp120", *Journal of Virology*, Vol. 87 No. 2, pp. 923–934.

Reed, L.J. and Muench, H. (1938), "A SIMPLE METHOD OF ESTIMATING FIFTY PER CENT ENDPOINTS", *American Journal of Epidemiology*, Vol. 27 No. 3, pp. 493–497.

Reeves, J.D., Gallo, S.A., Ahmad, N., Miamidian, J.L., Harvey, P.E., Sharron, M., Pöhlmann, S., et al. (2002), "Sensitivity of HIV-1 to entry inhibitors correlates with envelope/coreceptor affinity, receptor density, and fusion kinetics", *Proceedings of the National Academy of Sciences*, Vol. 99 No. 25, pp. 16249–16254.

Reeves, J.D., Lee, F.-H., Miamidian, J.L., Jabara, C.B., Juntilla, M.M. and Doms, R.W. (2005), "Enfuvirtide Resistance Mutations: Impact on Human Immunodeficiency Virus Envelope Function, Entry Inhibitor Sensitivity, and Virus Neutralization", *Journal of Virology*, Vol. 79 No. 8, pp. 4991–4999.

Repits, J. (2005), "Selection of human immunodeficiency virus type 1 R5 variants with augmented replicative capacity and reduced sensitivity to entry inhibitors during severe immunodeficiency", *Journal of General Virology*, Vol. 86 No. 10, pp. 2859–2869.

Requejo, H.I. (2006), "Worldwide molecular epidemiology of HIV", *Revista de Saude Publica*, Vol. 40 No. 2, pp. 331–345.

Rizzuto, C.D. (1998), "A Conserved HIV gp120 Glycoprotein Structure Involved in Chemokine Receptor Binding", *Science*, Vol. 280 No. 5371, pp. 1949–1953.

Robertson, D.L., Anderson, J.P., Bradac, J.A., Carr, J.K., Foley, B., Funkhouser, R.K., Gao, F., et al. (2000), "HIV-1 Nomenclature Proposal", *Science*, Vol. 288 No. 5463, pp. 55–55.

Rousseau, C.M., Learn, G.H., Bhattacharya, T., Nickle, D.C., Heckerman, D., Chetty, S., Brander, C., et al. (2007), "Extensive Intrasubtype Recombination in South African Human Immunodeficiency Virus Type 1 Subtype C Infections", *Journal of Virology*, Vol. 81 No. 9, pp. 4492–4500.

Sagar, M., Lavreys, L., Baeten, J.M., Richardson, B.A., Mandaliya, K., Chohan, B.H., Kreiss, J.K., et al. (2003), "Infection with Multiple Human Immunodeficiency Virus Type 1 Variants Is Associated with Faster Disease Progression", *Journal of Virology*, Vol. 77 No. 23, pp. 12921–12926.

Sagar, M., Lavreys, L., Baeten, J.M., Richardson, B.A., Mandaliya, K., Ndinya-Achola, J.O., Kreiss, J.K., et al. (2004), "Identification of modifiable factors that affect the genetic diversity of the transmitted HIV-1 population", *AIDS (London, England)*, Vol. 18 No. 4, pp. 615–619.

Sahni, A.K., Gupta, R.M., Nagendra, A., Nema, S.K., Rai, R. and Bhardwaj, J.R. (2007), "Identification of human immunodeficiency virus type-1 subtypes by heteroduplex mobility assay", *Medical Journal Armed Forces India*, Vol. 63 No. 3, pp. 249–252.

Sala, M., Zambruno, G., Vartanian, J.P., Marconi, A., Bertazzoni, U. and Wain-Hobson, S. (1994), "Spatial discontinuities in human immunodeficiency virus type 1

quasispecies derived from epidermal Langerhans cells of a patient with AIDS and evidence for double infection.", *Journal of Virology*, Vol. 68 No. 8, p. 5280.

Salazar-Gonzalez, J.F., Bailes, E., Pham, K.T., Salazar, M.G., Guffey, M.B., Keele, B.F., Derdeyn, C.A., et al. (2008), "Deciphering Human Immunodeficiency Virus Type 1 Transmission and Early Envelope Diversification by Single-Genome Amplification and Sequencing", *Journal of Virology*, Vol. 82 No. 8, pp. 3952–3970.

Salazar-Gonzalez, J.F., Salazar, M.G., Keele, B.F., Learn, G.H., Giorgi, E.E., Li, H., Decker, J.M., et al. (2009), "Genetic identity, biological phenotype, and evolutionary pathways of transmitted/founder viruses in acute and early HIV-1 infection", *The Journal of Experimental Medicine*, Vol. 206 No. 6, pp. 1273–1289.

Salazar-Gonzalez, J.F., Salazar, M.G., Learn, G.H., Fouda, G.G., Kang, H.H., Mahlokozera, T., Wilks, A.B., et al. (2011), "Origin and evolution of HIV-1 in breast milk determined by single-genome amplification and sequencing.", *Journal of Virology*, Vol. 85 No. 6, pp. 2751–63.

Sanborn, K.B., Somasundaran, M., Luzuriaga, K. and Leitner, T. (2015), "Recombination elevates the effective evolutionary rate and facilitates the establishment of HIV-1 infection in infants after mother-to-child transmission", *Retrovirology*, Vol. 12, p. 96

Maria Mercedes Santoro and Carlo Federico Perno, "HIV-1 Genetic Variability and Clinical Implications," *ISRN Microbiology*, vol. 2013, Article ID 481314, 20 pages, 2013. doi:10.1155/2013/481314

Santra, S., Liao, H.-X., Zhang, R., Muldoon, M., Watson, S., Fischer, W., Theiler, J., et al. (2010), "Mosaic vaccines elicit CD8+ T lymphocyte responses that confer enhanced immune coverage of diverse HIV strains in monkeys", *Nature Medicine*, Vol. 16 No. 3, pp. 324–328.

Sather, D.N., Carbonetti, S., Kehayia, J., Kraft, Z., Mikell, I., Scheid, J.F., Klein, F., et al. (2012), "Broadly Neutralizing Antibodies Developed by an HIV-Positive Elite Neutralizer Exact a Replication Fitness Cost on the Contemporaneous Virus", *Journal of Virology*, Vol. 86 No. 23, pp. 12676–12685.

Schnell, G., Joseph, S., Spudich, S., Price, R.W. and Swanstrom, R. (2011), "HIV-1 Replication in the Central Nervous System Occurs in Two Distinct Cell Types", edited by Cullen, B.R. *PLoS Pathogens*, Vol. 7 No. 10, p. e1002286.

Schramm, B., Penn, M.L., Speck, R.F., Chan, S.Y., De Clercq, E., Schols, D., Connor, R.I., et al. (2000), "Viral entry through CXCR4 is a pathogenic factor and therapeutic target in human immunodeficiency virus type 1 disease", *Journal of Virology*, Vol. 74 No. 1, pp. 184–192.

Selhorst, P., Combrinck, C., Ndabambi, N., Ismail, S.D., Abrahams, M.-R., Lacerda, M., Samsunder, N., et al. (2017), "REPLICATION CAPACITY OF VIRUSES FROM ACUTE INFECTION DRIVES HIV-1 DISEASE PROGRESSION", *Journal of Virology*, p. JVI.01806-16.

Shankarappa, R.A.J., Margolick, J.B., Gange, S.J., Rodrigo, A.G., Upchurch, D., Farzadegan, H., Gupta, P., et al. (1999), "Consistent viral evolutionary changes associated with the progression of human immunodeficiency virus type 1 infection", *Journal of Virology*, Vol. 73 No. 12, pp. 10489–10502.

Shen, R., Raska, M., Bimczok, D., Novak, J. and Smith, P.D. (2014), "HIV-1 Envelope Glycan Moieties Modulate HIV-1 Transmission", *Journal of Virology*, Vol. 88 No. 24, pp. 14258–14267.

Simon-Loriere, E., Galetto, R., Hamoudi, M., Archer, J., Lefeuvre, P., Martin, D.P., Robertson, D.L., et al. (2009), "Molecular Mechanisms of Recombination Restriction in the Envelope Gene of the Human Immunodeficiency Virus", edited by Holmes, E.C. *PLoS Pathogens*, Vol. 5 No. 5, p. e1000418.

Smith, D.M., Richman, D.D. and Little, S.J. (2005), "HIV superinfection", *Journal of Infectious Diseases*, Vol. 192 No. 3, pp. 438–444.

Smith, D.M., Wong, J.K., Hightower, G.K., Ignacio, C.C., Koelsch, K.K., Daar, E.S., Richman, D.D., et al. (2004), "Incidence of HIV superinfection following primary infection", *JAMA*, Vol. 292 No. 10, pp. 1177–1178.

Smith, D.M., Wong, J.K., Hightower, G.K., Ignacio, C.C., Koelsch, K.K., Petropoulos, C.J., Richman, D.D., et al. (2005), "HIV drug resistance acquired through superinfection", *Aids*, Vol. 19 No. 12, pp. 1251–1256.

Smyth, R.P., Davenport, M.P. and Mak, J. (2012), "The origin of genetic diversity in HIV-1", *Virus Research*, Vol. 169 No. 2, pp. 415–429.

Song, H., Hora, B., Bhattacharya, T., Goonetilleke, N., Liu, M.K.P., Wiehe, K., Li, H., et al. (2014), "Reversion and T Cell Escape Mutations Compensate the Fitness Loss of a CD8+ T Cell Escape Mutant in Their Cognate Transmitted/Founder Virus", edited by Vartanian, J.-P. *PLoS ONE*, Vol. 9 No. 7, p. e102734.

Song, H., Hora, B., Giorgi, E.E., Kumar, A., Cai, F., Bhattacharya, T., Perelson, A.S., et al. (2016), "Transmission of Multiple HIV-1 Subtype C Transmitted/founder Viruses into the Same Recipients Was not Determined by Modest Phenotypic Differences", *Scientific Reports*, Vol. 6, p. 38130.

Song, H., Pavlicek, J.W., Cai, F., Bhattacharya, T., Li, H., Iyer, S.S., Bar, K.J., et al. (2012), "Impact of immune escape mutations on HIV-1 fitness in the context of the cognate transmitted/founder genome", *Retrovirology*, Vol. 9 No. 1, p. 1.

Spira, A.I. and Ho, D.D. (1995), "Effect of different donor cells on human immunodeficiency virus type 1 replication and selection in vitro.", *Journal of Virology*, Vol. 69 No. 1, pp. 422–429.

Ssemwanga, D., Lyagoba, F., Ndembi, N., Mayanja, B.N., Larke, N., Wang, S., Baalwa, J., et al. (2011), "Multiple HIV-1 infections with evidence of recombination in heterosexual partnerships in a low risk Rural Clinical Cohort in Uganda", *Virology*, Vol. 411 No. 1, pp. 113–131.

Staropoli, I., Chanel, C., Girard, M. and Altmeyer, R. (2000), "Processing, Stability, and Receptor Binding Properties of Oligomeric Envelope Glycoprotein from a Primary HIV-1 Isolate", *Journal of Biological Chemistry*, Vol. 275 No. 45, pp. 35137–35145.

Sterjovski, J., Churchill, M.J., Ellett, A., Gray, L.R., Roche, M.J., Dunfee, R.L., Purcell, D.F., et al. (2007), "Asn 362 in gp120 contributes to enhanced fusogenicity by CCR5-restricted HIV-1 envelope glycoprotein variants from patients with AIDS", *Retrovirology*, Vol. 4 No. 1, p. 89.

Stieh, D.J., King, D.F., Klein, K., Aldon, Y., McKay, P.F. and Shattock, R.J. (2015), "Discrete partitioning of HIV-1 Env forms revealed by viral capture", *Retrovirology*, Vol. 12 No. 1, p. 81.

Sucupira, M.C.A., Sanabani, S., Cortes, R.M., Giret, M.T.M., Tomiyama, H., Sauer, M.M., Sabino, E.C., et al. (2012), "Faster HIV-1 Disease Progression among Brazilian Individuals Recently Infected with CXCR4-Utilizing Strains", edited by Poehlmann, S. *PLoS ONE*, Vol. 7 No. 1, p. e30292.

Sunshine, J.E., Larsen, B.B., Maust, B., Casey, E., Deng, W., Chen, L., Westfall, D.H., et al. (2015), "Fitness-Balanced Escape Determines Resolution of Dynamic Founder Virus Escape Processes in HIV-1 Infection", edited by Silvestri, G. *Journal of Virology*, Vol. 89 No. 20, pp. 10303–10318.

Tamura, K., Peterson, D., Peterson, N., Stecher, G., Nei, M. and Kumar, S. (2011), "MEGA5: Molecular Evolutionary Genetics Analysis Using Maximum Likelihood, Evolutionary Distance, and Maximum Parsimony Methods", *Molecular Biology and Evolution*, Vol. 28 No. 10, pp. 2731–2739.

Taylor, J.R., Kimbrell, K.C., Scoggins, R., Delaney, M., Wu, L. and Camerini, D. (2001), "Expression and Function of Chemokine Receptors on Human Thymocytes: Implications for Infection by Human Immunodeficiency Virus Type 1", *Journal of Virology*, Vol. 75 No. 18, pp. 8752–8760.

Tebit, D.M., Nankya, I., Arts, E.J. and Gao, Y. (2007), "HIV diversity, recombination and disease progression: how does fitness 'fit' into the puzzle", *AIDS Rev*, Vol. 9 No. 2, pp. 75–87.

Tomaras, G.D., Yates, N.L., Liu, P., Qin, L., Fouda, G.G., Chavez, L.L., Decamp, A.C., et al. (2008), "Initial B-Cell Responses to Transmitted Human Immunodeficiency Virus Type 1: Virion-Binding Immunoglobulin M (IgM) and IgG Antibodies Followed by Plasma Anti-gp41 Antibodies with Ineffective Control of Initial Viremia", *Journal of Virology*, Vol. 82 No. 24, pp. 12449–12463.

Townsley, S., Li, Y., Kozyrev, Y., Cleveland, B. and Hu, S.-L. (2016), "Conserved Role of an N-Linked Glycan on the Surface Antigen of Human Immunodeficiency Virus Type 1 Modulating Virus Sensitivity to Broadly Neutralizing Antibodies against the Receptor and Coreceptor Binding Sites", edited by Silvestri, G.*Journal of Virology*, Vol. 90 No. 2, pp. 829–841.

Trkola, A., Kuster, H., Leemann, C., Ruprecht, C., Joos, B., Telenti, A., Hirschel, B., et al. (2003), "Human Immunodeficiency Virus Type 1 Fitness Is a Determining Factor in

Viral Rebound and Set Point in Chronic Infection", *Journal of Virology*, Vol. 77 No. 24, pp. 13146–13155.

Trkola, A., Pomales, A.B., Yuan, H., Korber, B., Maddon, P.J., Allaway, G.P., Katinger, H., et al. (1995), "Cross-clade neutralization of primary isolates of human immunodeficiency virus type 1 by human monoclonal antibodies and tetrameric CD4-IgG.", *Journal of Virology*, Vol. 69 No. 11, pp. 6609–6617.

Troyer, R.M., Collins, K.R., Abraha, A., Fraundorf, E., Moore, D.M., Krizan, R.W., Toossi, Z., et al. (2005), "Changes in Human Immunodeficiency Virus Type 1 Fitness and Genetic Diversity during Disease Progression", *Journal of Virology*, Vol. 79 No. 14, pp. 9006–9018.

Troyer, R.M., McNevin, J., Liu, Y., Zhang, S.C., Krizan, R.W., Abraha, A., Tebit, D.M., et al. (2009), "Variable Fitness Impact of HIV-1 Escape Mutations to Cytotoxic T Lymphocyte (CTL) Response", edited by Walker, C.M. *PLoS Pathogens*, Vol. 5 No. 4, p. e1000365.

Tsibris, A.M.N., Sagar, M., Gulick, R.M., Su, Z., Hughes, M., Greaves, W., Subramanian, M., et al. (2008), "In Vivo Emergence of Vicriviroc Resistance in a Human Immunodeficiency Virus Type 1 Subtype C-Infected Subject", *Journal of Virology*, Vol. 82 No. 16, pp. 8210–8214.

Tsui, R., Herring, B.L., Barbour, J.D., Grant, R.M., Bacchetti, P., Kral, A., Edlin, B.R., et al. (2004), "Human Immunodeficiency Virus Type 1 Superinfection Was Not Detected following 215 Years of Injection Drug User Exposure", *Journal of Virology*, Vol. 78 No. 1, pp. 94–103.

Tuttle, D.L., Anders, C.B., Aquino-De Jesus, M.J., Poole, P.P., Lamers, S.L., Briggs, D.R., Pomeroy, S.M., et al. (2002), "Increased replication of non-syncytium-inducing HIV type 1 isolates in monocyte-derived macrophages is linked to advanced disease in infected children", *AIDS Research and Human Retroviruses*, Vol. 18 No. 5, pp. 353–362.

van den Kerkhof, T.L.G.M., de Taeye, S.W., Boeser-Nunnink, B.D., Burton, D.R., Kootstra, N.A., Schuitemaker, H., Sanders, R.W., et al. (2016), "HIV-1 escapes from N332-directed antibody neutralization in an elite neutralizer by envelope glycoprotein elongation and introduction of unusual disulfide bonds", *Retrovirology*, Vol. 13 No. 1, p. 48.

van der Kuyl, A.C. and Cornelissen, M. (2007), "Identifying HIV-1 dual infections.", *Retrovirology*, Vol. 4, p. 67.

van Gils, M.J., Bunnik, E.M., Burger, J.A., Jacob, Y., Schweighardt, B., Wrin, T. and Schuitemaker, H. (2010), "Rapid Escape from Preserved Cross-Reactive Neutralizing Humoral Immunity without Loss of Viral Fitness in HIV-1-Infected Progressors and Long-Term Nonprogressors", *Journal of Virology*, Vol. 84 No. 7, pp. 3576–3585.

van Loggerenberg, F., Mlisana, K., Williamson, C., Auld, S.C., Morris, L., Gray, C.M., Abdool Karim, Q., et al. (2008), "Establishing a Cohort at High Risk of HIV Infection

in South Africa: Challenges and Experiences of the CAPRISA 002 Acute Infection Study", edited by Ugarte-Gil, C.A. *PLoS ONE*, Vol. 3 No. 4, p. e1954.

van Rij, R.P., Hazenberg, M.D., van Benthem, B.H.B., Otto, S.A., Prins, M., Miedema, F. and Schuitemaker, H. (2003), "Early Viral Load and CD4+ T Cell Count, But Not Percentage of CCR5+ or CXCR4+ CD4+ T Cells, Are Associated with R5-to-X4 HIV Type 1 Virus Evolution", *AIDS Research and Human Retroviruses*, Vol. 19 No. 5, pp. 389–398.

Varmus, H. (1988), "Retroviruses", *Science (New York, N.Y.)*, Vol. 240 No. 4858, pp. 1427–1435.

Vaux, D.L. (2012), "Research methods: Know when your numbers are significant", *Nature*, Vol. 492 No. 7428, pp. 180–181.

Veronese, F.D., Copeland, T.D., Oroszlan, S., Gallo, R.C. and Sarngadharan, M.G. (1988), "Biochemical and immunological analysis of human immunodeficiency virus gag gene products p17 and p24.", *Journal of Virology*, Vol. 62 No. 3, pp. 795–801.

Wagner, G.A., Pacold, M.E., Kosakovsky Pond, S.L., Caballero, G., Chaillon, A., Rudolph, A.E., Morris, S.R., et al. (2014), "Incidence and Prevalence of Intrasubtype HIV-1 Dual Infection in At-Risk Men in the United States", *The Journal of Infectious Diseases*, Vol. 209 No. 7, pp. 1032–1038.

Walter, B.L., Wehrly, K., Swanstrom, R., Platt, E., Kabat, D. and Chesebro, B. (2005), "Role of Low CD4 Levels in the Influence of Human Immunodeficiency Virus Type 1 Envelope V1 and V2 Regions on Entry and Spread in Macrophages", *Journal of Virology*, Vol. 79 No. 8, pp. 4828–4837.

Wang, W., Nie, J., Prochnow, C., Truong, C., Jia, Z., Wang, S., Chen, X.S., et al. (2013), "A systematic study of the N-glycosylation sites of HIV-1 envelope protein on infectivity and antibody-mediated neutralization.", *Retrovirology*, Vol. 10, p. 14.

Wang, Y.E., Li, B., Carlson, J.M., Streeck, H., Gladden, A.D., Goodman, R., Schneidewind, A., et al. (2009), "Protective HLA Class I Alleles That Restrict Acute-Phase CD8+ T-Cell Responses Are Associated with Viral Escape Mutations Located in Highly Conserved Regions of Human Immunodeficiency Virus Type 1", *Journal of Virology*, Vol. 83 No. 4, pp. 1845–1855.

Weber, J., Vazquez, A.C., Winner, D., Rose, J.D., Wylie, D., Rhea, A.M., Henry, K., et al. (2011), "Novel Method for Simultaneous Quantification of Phenotypic Resistance to Maturation, Protease, Reverse Transcriptase, and Integrase HIV Inhibitors Based on 3′Gag(p2/p7/p1/p6)/PR/RT/INT-Recombinant Viruses: a Useful Tool in the Multitarget Era of Antiretroviral Therapy▿", *Antimicrobial Agents and Chemotherapy*, Vol. 55 No. 8, pp. 3729–3742.

Wei, X., Decker, J.M., Liu, H., Zhang, Z., Arani, R.B., Kilby, J.M., Saag, M.S., et al. (2002), "Emergence of Resistant Human Immunodeficiency Virus Type 1 in Patients Receiving Fusion Inhibitor (T-20) Monotherapy", *Antimicrobial Agents and Chemotherapy*, Vol. 46 No. 6, pp. 1896–1905.

Wei, X., Decker, J.M., Wang, S., Hui, H., Kappes, J.C., Wu, X., Salazar-Gonzalez, J.F., et al. (2003), "Antibody neutralization and escape by HIV-1", *Nature*, Vol. 422 No. 6929, pp. 307–312.

Wei, X., Ghosh, S.K., Taylor, M.E., Johnson, V.A., Emini, E.A., Deutsch, P., Lifson, J.D., et al. (1995), "Viral dynamics in human immunodeficiency virus type 1 infection", *Nature*, Vol. 373 No. 6510, pp. 117–122.

Wilen, C.B., Tilton, J.C. and Doms, R.W. (2012a), "Molecular Mechanisms of HIV Entry", in Rossmann, M.G. and Rao, V.B. (Eds.), *Viral Molecular Machines*, Vol. 726, Springer US, Boston, MA, pp. 223–242.

Wilen, C.B., Tilton, J.C. and Doms, R.W. (2012b), "HIV: Cell Binding and Entry", *Cold Spring Harbor Perspectives in Medicine*, Vol. 2 No. 8, pp. a006866–a006866.

Williamson, C. and Swanstrom, R. (2015), "HIV-1 replication capacity: Setting the pace of disease", *Proceedings of the National Academy of Sciences*, p. 201502208.

Wood, P. and Elliott, T. (1998), "Glycan-regulated Antigen Processing of a Protein in the Endoplasmic Reticulum Can Uncover Cryptic Cytotoxic T Cell Epitopes", *The Journal of Experimental Medicine*, Vol. 188 No. 4, pp. 773–778.

Woodman, Z., Mlisana, K., Treurnicht, F., Abrahams, M.-R., Thebus, R., Karim, S.A. and Williamson, C. (2011), "Short Communication Decreased Incidence of Dual Infections in South African Subtype C-Infected Women Compared to a Cohort Ten Years Earlier", *AIDS Research and Human Retroviruses*, Vol. 27 No. 11, pp. 1167–72.

Wooley, D.P., Smith, R.A., Czajak, S. and Desrosiers, R.C. (1997), "Direct demonstration of retroviral recombination in a rhesus monkey.", *Journal of Virology*, Vol. 71 No. 12, pp. 9650–9653.

Wu, Y. (2004), "HIV-1 gene expression: lessons from provirus and non-integrated DNA", *Retrovirology*, Vol. 1 No. 1, p. 13.

Wyatt, R. (1998), "The HIV-1 Envelope Glycoproteins: Fusogens, Antigens, and Immunogens", *Science*, Vol. 280 No. 5371, pp. 1884–1888.

Wyma, D.J., Jiang, J., Shi, J., Zhou, J., Lineberger, J.E., Miller, M.D. and Aiken, C. (2004), "Coupling of Human Immunodeficiency Virus Type 1 Fusion to Virion Maturation: a Novel Role of the gp41 Cytoplasmic Tail", *Journal of Virology*, Vol. 78 No. 7, pp. 3429–3435.

Wyma, D.J., Kotov, A. and Aiken, C. (2000), "Evidence for a stable interaction of gp41 with Pr55Gag in immature human immunodeficiency virus type 1 particles", *Journal of Virology*, Vol. 74 No. 20, pp. 9381–9387.

Wyss, S., Dimitrov, A.S., Baribaud, F., Edwards, T.G., Blumenthal, R. and Hoxie, J.A. (2005), "Regulation of Human Immunodeficiency Virus Type 1 Envelope Glycoprotein Fusion by a Membrane-Interactive Domain in the gp41 Cytoplasmic Tail", *Journal of Virology*, Vol. 79 No. 19, pp. 12231–12241.

Yerly, S., Jost, S., Monnat, M., Telenti, A., Cavassini, M., Chave, J.-P., Kaiser, L., et al. (2004), "HIV-1 co/super-infection in intravenous drug users", *AIDS*, Vol. 18 No. 10, pp. 1413–1421.

Yi, H.A., Diaz-Rohrer, B., Saminathan, P. and Jacobs, A. (2015), "The Membrane Proximal External Regions of gp41 from HIV-1 Strains HXB2 and JRFL Have Different Sensitivities to Alanine Mutation", *Biochemistry*, Vol. 54 No. 8, pp. 1681–1693.

Ying, R., Granich, R.M., Gupta, S. and Williams, B.G. (2016), "CD4 Cell Count: Declining Value for Antiretroviral Therapy Eligibility", *Clinical Infectious Diseases: An Official Publication of the Infectious Diseases Society of America*, Vol. 62 No. 8, pp. 1022–1028.

Yu, X., Yuan, X., Mclane, M.F., Lee, T.H. and Essex, M. (1993), "Mutations in the cytoplasmic domain of human immunodeficiency virus type 1 transmembrane protein impair the incorporation of Env proteins into mature virions.", *Journal of Virology*, Vol. 67 No. 1, pp. 213–221.

Yue, L., Shang, L. and Hunter, E. (2009), "Truncation of the Membrane-Spanning Domain of Human Immunodeficiency Virus Type 1 Envelope Glycoprotein Defines Elements Required for Fusion, Incorporation, and Infectivity", *Journal of Virology*, Vol. 83 No. 22, pp. 11588–11598.

Yuste, E., Reeves, J.D., Doms, R.W. and Desrosiers, R.C. (2004), "Modulation of Env Content in Virions of Simian Immunodeficiency Virus: Correlation with Cell Surface Expression and Virion Infectivity", *Journal of Virology*, Vol. 78 No. 13, pp. 6775–6785.

Zaunders, J. and van Bockel, D. (2013), "Innate and Adaptive Immunity in Long-Term Non-Progression in HIV Disease", *Frontiers in Immunology*, Vol. 4, p. 95.

Zhang, Z., Li, S., Gu, Y. and Xia, N. (2016), "Antiviral Therapy by HIV-1 Broadly Neutralizing and Inhibitory Antibodies", *International Journal of Molecular Sciences*, Vol. 17 No. 11, p. 1901.

Zhu, T., Wang, N., Carr, A., Wolinsky, S. and Ho, D.D. (1995), "Evidence for coinfection by multiple strains of human immunodeficiency virus type 1 subtype B in an acute seroconvertor.", *Journal of Virology*, Vol. 69 No. 2, pp. 1324–1327.

APPENDIX A.

Primer Sequences and PCR Conditions

Primers used for *envelope* PCR amplification

Primers Name	Sequence 5'- 3'	HXB2 location
Env N	TTGCCAATCAAGGAAGTAGCCTTGTGT	9145←9171
Env 1M	TAGCCCTTCCAGTCCCCCCTTTTCTTTTA	9068←9096
Env 1A-RX	CACCGGCTTAGGCATCTCCTATAGCAGGAAGAA	5950→5982
Env IF	AGA AAG AGC AGA AGA CAG TGG CAA TGA	6202→6228
Env IR	TTT TGA CCA CTT GCC ACC CAT	8797←8817
Rev 15	CTGCCATTTAACAGCAGTTGAGGTGA	6990←7012
T7	TAATACGACTCACTATAGGG	Not applicable
Env-y F	AATGTCAGCACAGTACAATGTACACATGG	6944→6973
Env-y R	GGAGCTGTTGATCCTTTAGGTATCTTTC	7973←8001

Primers used for Site-directed mutagenesis (SDM)

Primers Name	Sequence 5'- 3'
84c4/ 84c1sdm(F)	AAAACAGACTGGAATAAAACTCTAGAAAAGGTAAAGGAA
84c4/ 84c1sdm (R)	TTCCTTTACCTTTTCTAGAGTTTTATTCCAGTCTGTTTT
84c4-sdm (F)	AAAACAGACTGGGAAAAAACTCTCGAGAGGGTAAAGGAA
84c4-sdm (R)	TTCCTTTACCCTCTCGAGAGTTTTTTCCCAGTCTGTTTT
137c10sdm (F)	ATAATAGGAGATATCAGACAAGCACATTGTAGCGTTAGTGGATGGAAT
137c10sdm (R)	ATTCCATCCACTAACGCTACAATGTGCTTGTCTGATATCTCCTATTAT

PCR mixture and cycling condition for 2nd round *env* amplification and SDM

Reagent	Concentration (Final)	Temperature (°C)	Time (min)	No. of repeats
Buffer + $MgCl_2$	1 X	94	5	-
dNTP mix	0.2 mM	94	0.30	
First primer (F)	0.4 mM	55*	0.30	45 times
Second primer (R)	0.4 mM	72	4	
Amplicon	1 µl	72	10	-
Phusion Hot Start Polymerase	2.5 U			
Total (made up with dH20)	50 µL			

* gradient PCR (55-62 °C) annealing temperature was used.

'A' tailing of the 2nd round env PCR

Reagent	Concentration (Final)	Temperature (°C)	Time (min)	No. of repeats
Buffer	1 X	72	10	-
MgCl2	2.5 mM			
dATPs	0.2 mM			
PCR product	18 µL			
Super Therm Taq polymerase	1.5 µL			
Total (made up with dH20)	50 µL			

PCR mixture and cycling condition for screening the correct *env* cloning

Reagent	Concentration (Final)
Buffer	1 X
$MgCl_2$	2mM
dNTP mix	0.2 mM
Primer (F)	0.4 mM
Primer (R)	0.4 mM
Amplicon	1 µl
Supertherm Taq	2.5 U
Total (made up with dH20)	50 µL

Temperature (°C)	Time (min)	No. of repeats
94	5	-
94	0.45	35 times
57	0.45	
72	4	
72	10	-

APPENDIX B.

Alignments and Phylogenetic Trees

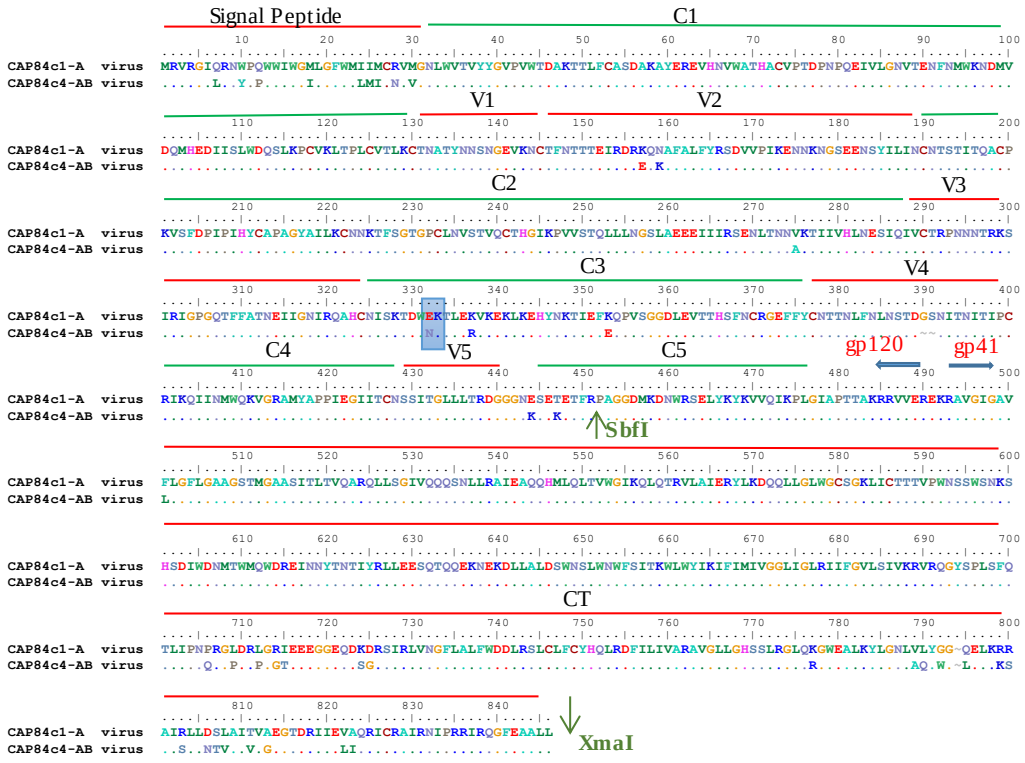


Figure B.1. CAP84 Env sequence alignment. The Env sequences of two clones of CAP84 were aligned using BioEdit program. CAP84c1 representing virus A at enrolment and CAP84c4 representing AB virus at 12 mpi. Different Env regions are indicated. Restriction enzyme sites used for digesting the two clones for generating the chimeras are indicated as well. A potential N-glycan site at 339 of CAP84c4 was detected using N-glycosite (www.HIV.lanl.gov) and highlighted with blue.

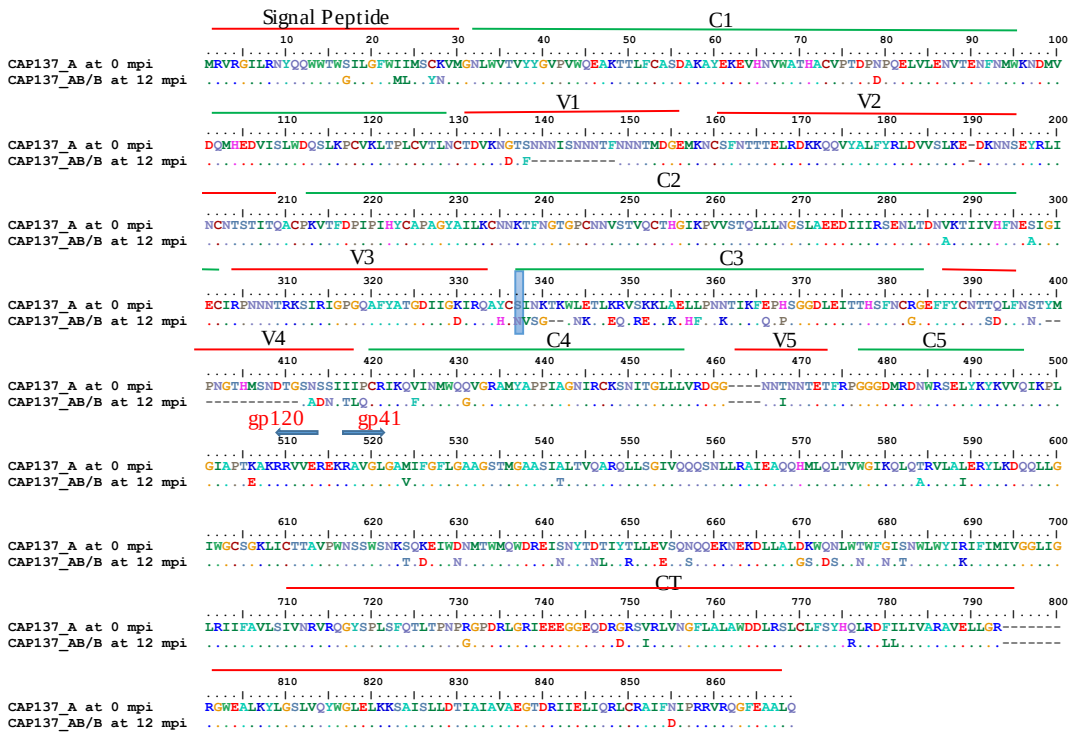


Figure B.2. CAP137 Env sequence alignment. The Env sequences of three clones of CAP137 were aligned using BioEdit program. Viruses A, and AB/B at 12 mpi are represented by 137c1 and CAP137c10, respectively. Different Env regions are indicated. A potential N-glycan site was detected using N-glycosite (www.HIV.lanl.gov) and highlighted with blue.

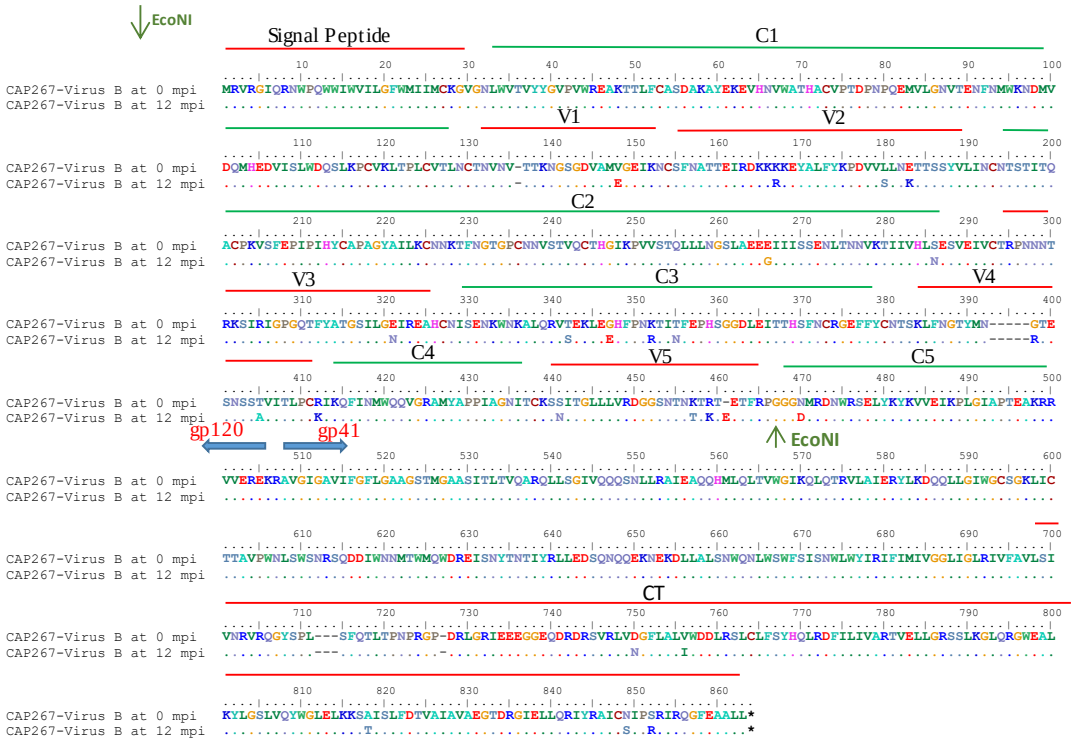


Figure B.3. CAP267 Env sequence alignment. The Env sequences of two clones of CAP267 were aligned using BioEdit program. CAP267c2 representing virus B at enrolment and 267c6 representing B virus at 12 mpi. Different Env regions were indicated. The Restriction enzyme sites used for digesting the two clones for generating the chimeras are indicated as well.

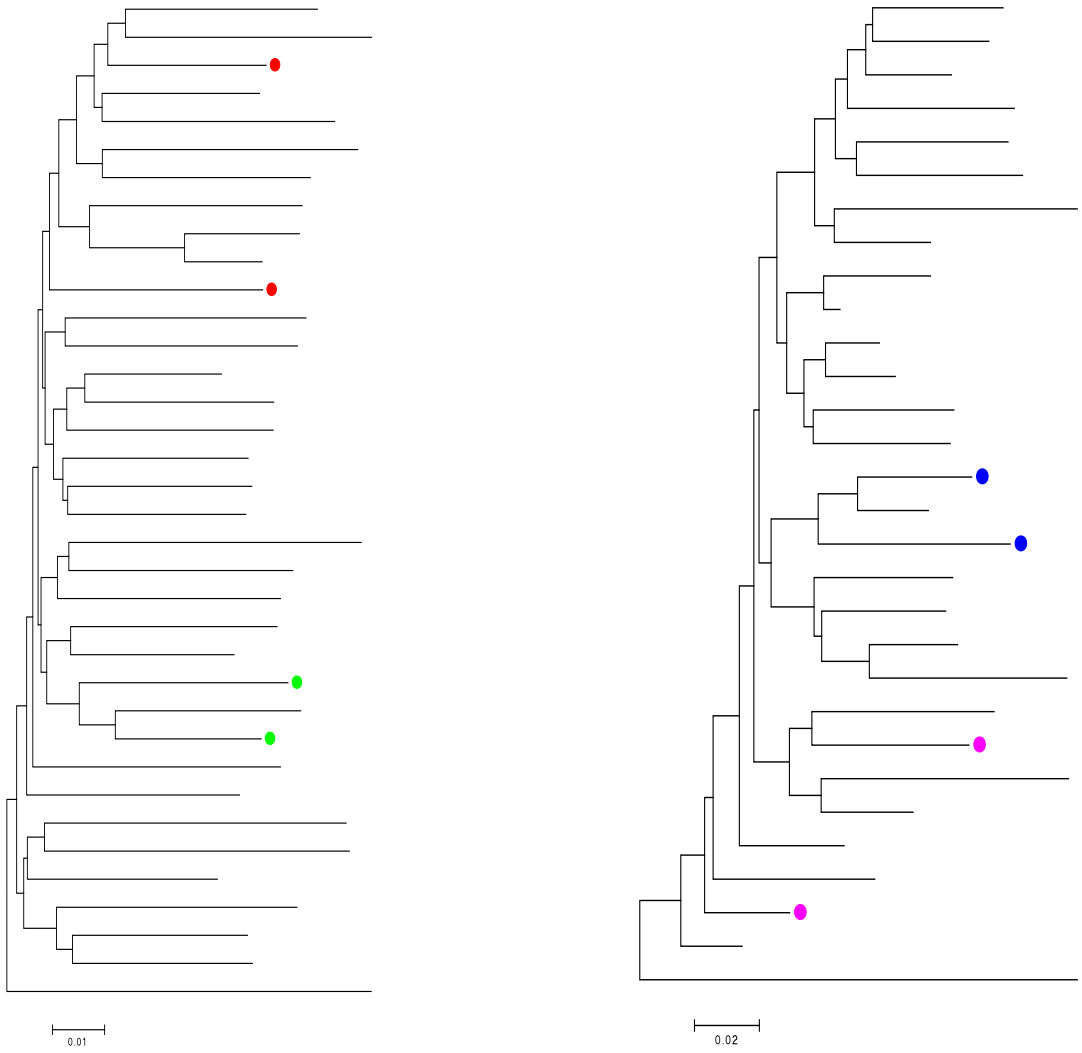


Figure B.4. Confirmation of infection with two phylogenetically distinct variants. Neighbour Joining trees of sub-genomic *env* sequences of the master sequences with randomly selected subtype C *env* reference sequences from South Africa. The two masters were separated by unlinked viruses. A) Infecting viruses from CAP37 and CAP84, labelled with red colour dots and green dots, respectively, and B) from CAP137 and CAP267 labelled with pink and blue respectively.

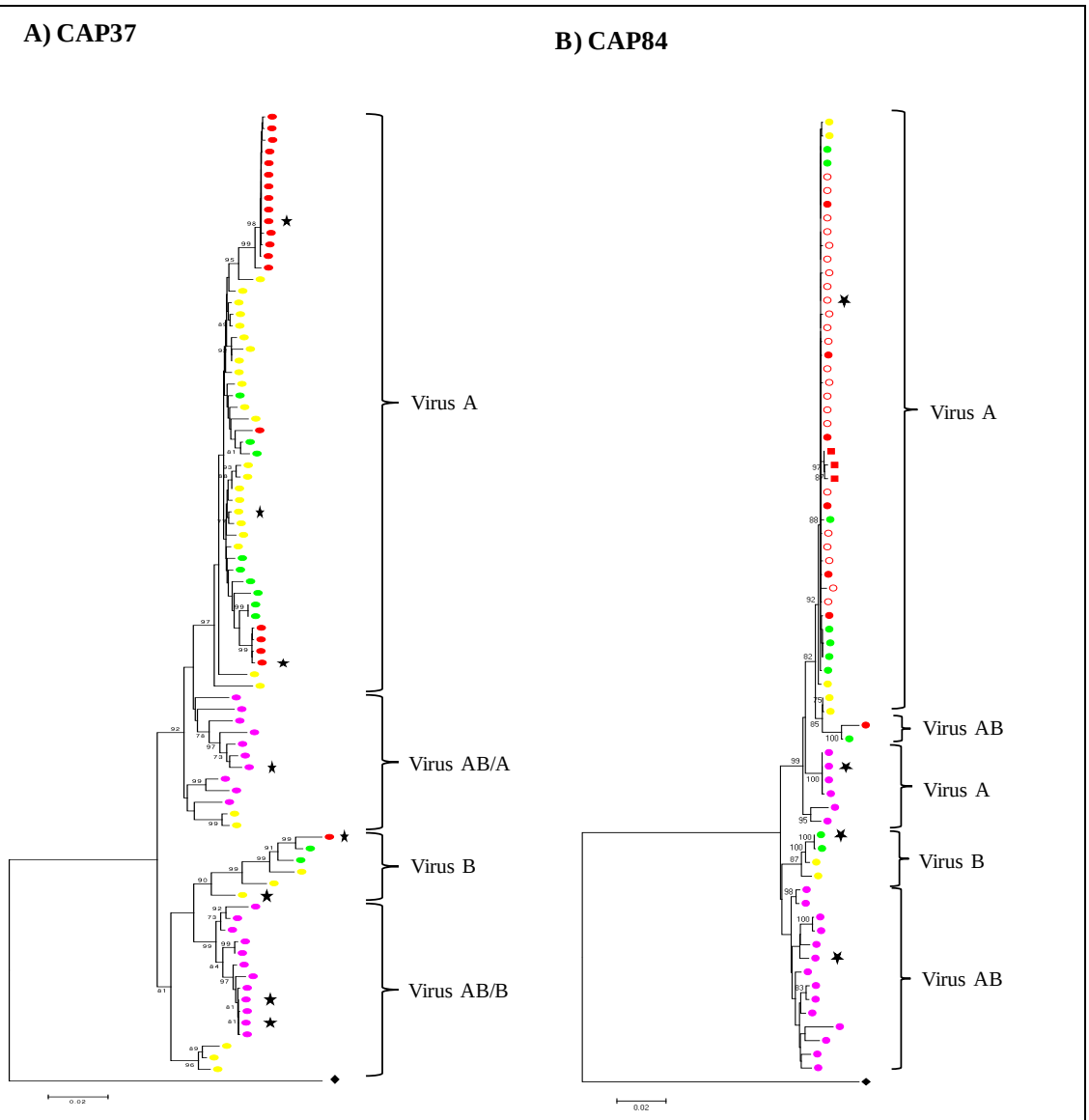


Figure B.5. Neighbour joining tree of SGA-*env* sequences of CAP37 and CAP84. The full-length SGA-derived *env* sequences of A) CAP37, B) CAP84, from 0, 3, 6 and 12 mpi were aligned using BioEdit. Neighbour-joining trees (Mega 5.0) were used to classified A (red) and B (blue) master sequences with the highest maximum DNA distance and AB variants (in purple) with equidistant maximum DNA distance from both masters. CAP37 sequences from 0 mpi, 3, 6, and 12 mpi are indicated with red, green, yellow and purple circle, respectively. CAP84 sequences from 0.25 mpi, 0.5 mpi, 1 mpi, 3 mpi, 6 mpi and 12 mpi are indicated with empty circle, red circle, filled square, green, yellow and purple circles, respectively. Boot strap values lower than 70 % are not shown.

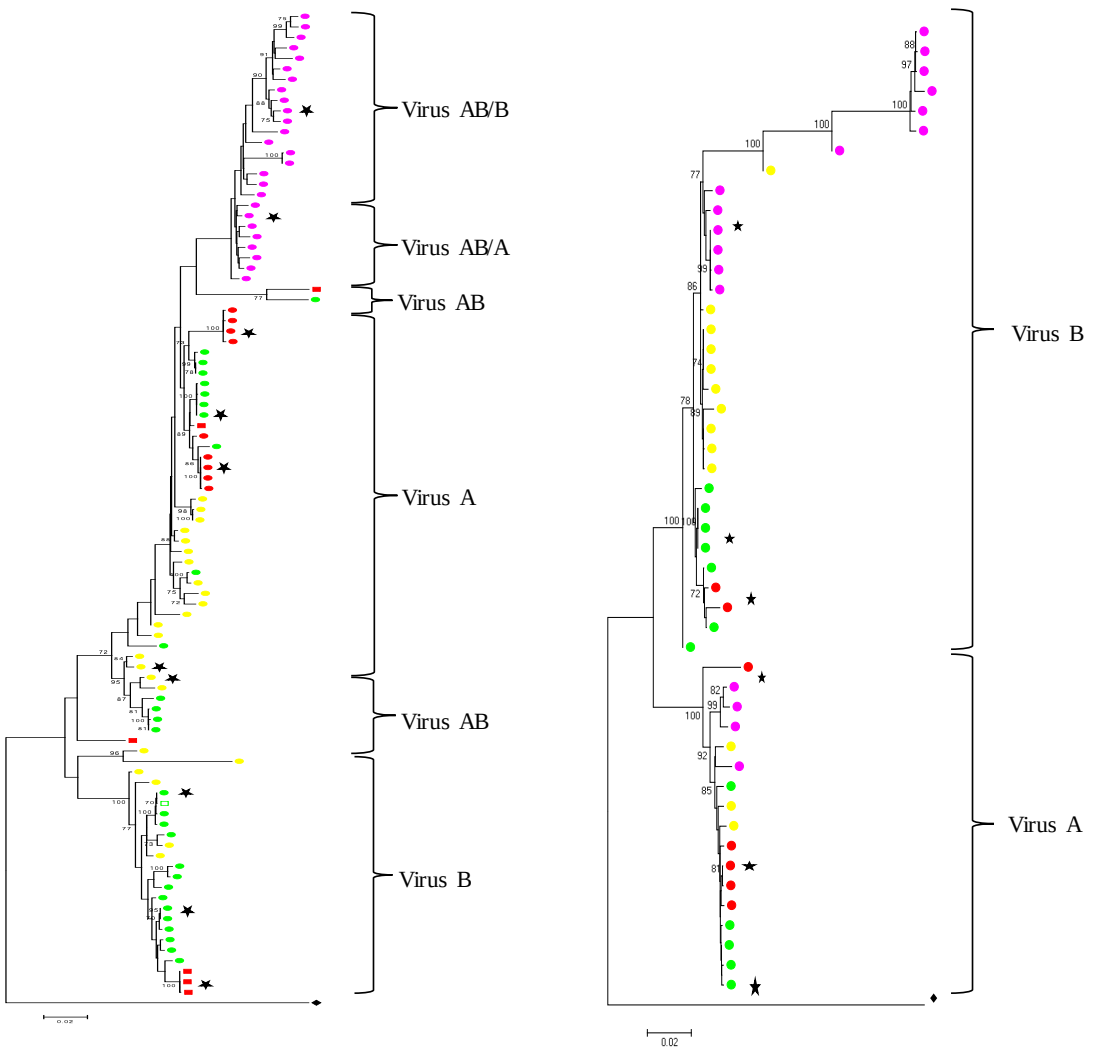


Figure B.6. Neighbour joining tree of SGA-*env* sequences of CAP137 and CAP267. The full-length SGA-derived *env* sequences of A) CAP137, B) CAP267, from 0, 3, 6 and 12 mpi were aligned using BioEdit. Neighbour-joining trees (Mega 5.0) were used to classified A (red) and B (blue) master sequences with the highest maximum DNA distance and AB variants (in purple) with equidistant maximum DNA distance from both masters. Sequences from 0 mpi, 3, 6, and 12 mpi are indicated with red, green, yellow and purple circle, respectively, and sequences from 2 mpi of CAP137 are indicted with red square. Boot strap values lower than 70 % are not shown.

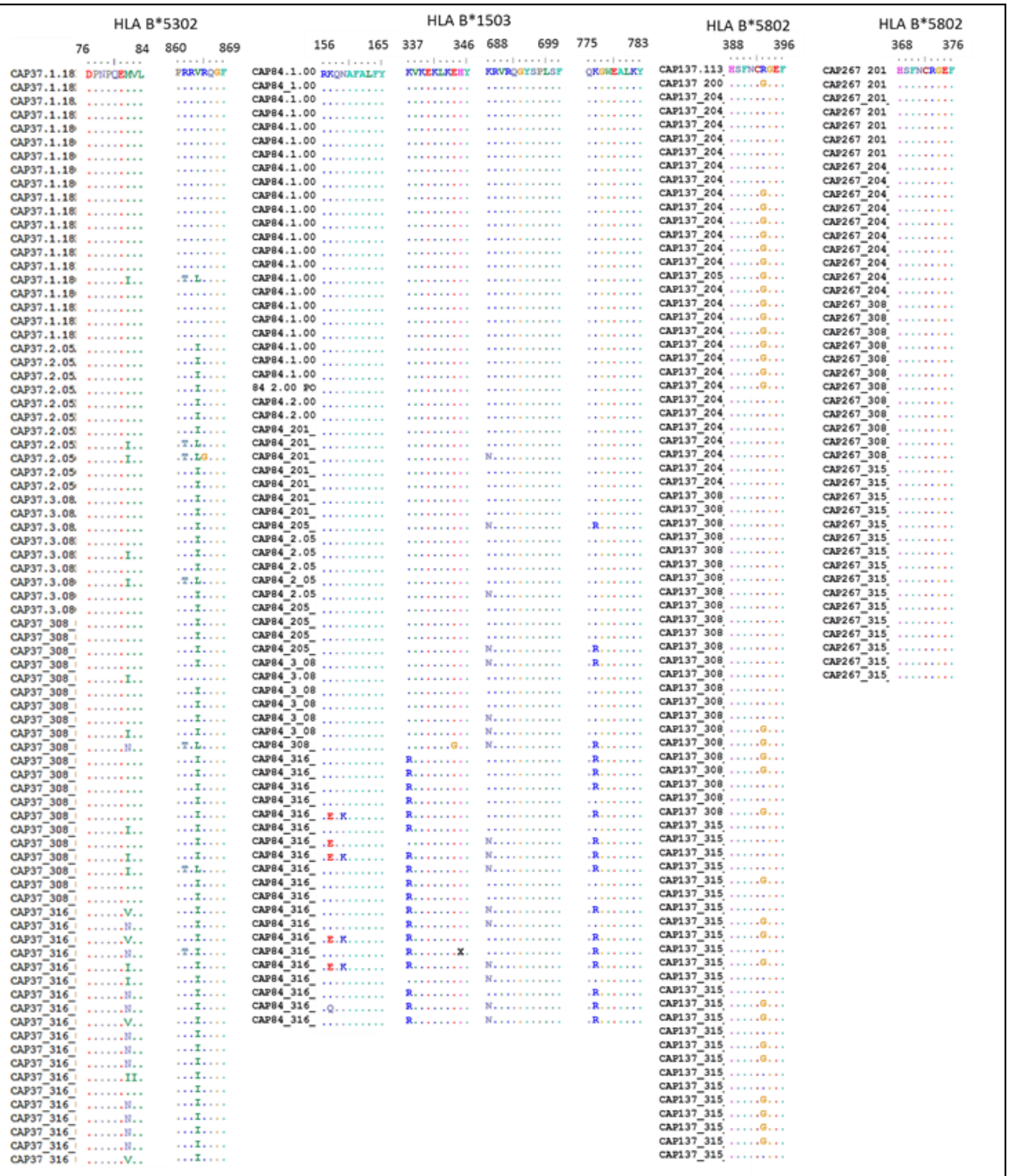


Figure B.7. Epitopes targeted by HLA B* alleles previously linked to disease progression. Sequences were scanned for epitopes recognised by HLA B* alleles specific to the dual infected individuals (Abrahams et al., 2013) that were shown to be linked to disease progression (Kiepiela et al., 2007; Prentice and Tang, 2012) using Motif Scan (www.lanl.gov). The epitopes are representatives to illustrate the fixation of putative CTL escape mutations.

APPENDIX C.

Media and Solutions

C.1 Media

Luria- Betani (LB) Broth

Mix the following:

Tryptone	10 g
NaCl	10 g
Yeast extract	5 g
ddH2O up to	1 L

Autoclave and store at room temperature.

If antibiotic media is required, add antibiotic Carbinocilline after autoclaving and prior to use immediately.

Luria Betani (LB) Plates

Mix the following:

Tryptone	10 g
NaCl	10 g
Yeast extract	5 g
Agar	15 g
ddH2O up to	1 L

Autoclave and cool to 55 °C in water bath. Add the required amount of Carbinocilline (100 mg/ml) or ampicillin in (100 μg/mL final concentration). Mix and pour into plates. Allow the plates to dry at room temperature before storing at 4 °C.

Yeast extract peptone dextrose (YEPD) media

Mix the following:

Yeast Extract	10 g
Peptone	20 g
Dextrose (D-glucose)	20 g
ddH2O up to	1 L

Autoclave and store at room temperature.

Yeast Amino Acid Dropout plates + FOA

Set the water bath to 55 °C

1- Prepare the following (use a 500 bottle):

Agar	20 g
ddH2O up to	500 ml

Autoclave. When agar is finished autoclaving, allow the agar to cool in the water bath for ten minutes.

2- Mix the following (use another 500-mL bottle):

Yeast Nitrogen Base (YNB)	6.7 g
Dextrose (D-glucose)	20 g
Complete Supplement Mixture minus Leucine (CSM-Leu)	0.69 g
FOA	1 g
ddH2O up to	500 ml

Stir until completely dissolved, Filter sterilise and Put in water bath to warm. Add the filter sterilised 500 ml mixture from above (the one that should be in the water bath as well in number 2) into the agar mixture and mix it slowly prior to pour into the plates. Allow the plates to cool at room temperature in the dark due to the possible inactivation of 5 - FOA by light.

(-Leu) Agar plates

- Set the water bath to 55 °C

1- Prepare the following (use a 500 bottle):

Agar	20 g
ddH2O up to	500 mL

Autoclave. When agar is finished autoclaving, allow the agar to cool in the water bath for ten minutes.

2- Mix the following:

Yeast Nitrogen Base (YNB)	6.7 g
Dextrose (D-glucose)	20 g
Complete Supplement Mixture minus Leucine (CSM-Leu)	0.69 g
ddH2O up to	350 ml

Stir until completely dissolved, filter sterilise and put in water bath to warm. Add the filter sterilised mixture from above (the one that should be in the water bath as well in number 2) into the agar mixture and mixed slowly prior to pour into the plates.

(-Leu) media yeast drop out

Mix the following:

Yeast Nitrogen Base (YNB)	6.7 g
Dextrose (D-glucose)	20 g
Complete Supplement Mixture minus Leucine (CSM-Leu)	0.69 g
ddH2O up to	1 L

Mix and autoclave

C.2 Solutions

Solutions for Agarose gel

50 X TAE (Tris-acetate buffer)

Tris	242 g
Glacial acetic acid	57.1 ml
EDTA (0.5 M)	100 ml
ddH2O up to	1 L

Ethidium Bromide (EtBr) Solution

EtBr (Sigma)	0.1 g
ddH2O up to	10 ml

Shake well to dissolve. Do not autoclave. Store in a dark bottle at room temperature. This is a powerful carcinogene, gloves should be worn at all times when handling this solution.

0.8 % agarose

Weigh 0.8 g agarose and add to 100 ml Schoot bottle. Add 50 ml 1X TAE. Boil in microwave and leave it to cool.

Solution for Pseudovirion Western Blotting

RIPA lysis buffer

Tris buffer, pH 7.5	10 mM
Sodium chloride	150 mM
Na2EDTA ph 8.0	2 mM
Triton X-100 detergent	1 %

1 tablet of complete protease inhibitor cocktail [Santa Cruz Biotechnology®] per 50 mL buffer solution). Aliquot into 1 ml and store at -20 °C.

10X SDS running buffer

Trizma (250mM)	30 g
Glycine (2M)	114 g
SDS 20 %	50 ml
ddH2O up to	1 L

10X TBS

Tris (200mM)	24.2 g
NaCl	80 g

Add 800 ml ddH2O. pH to 7.6 with HCl. Make up to 1 liter with ddH2O

4X SDS loading buffer

Tris (1M) pH 6.8	2.5 ml
SDS (20 %)	1 ml
β-mercaptoethanol	2 ml
Glycerol	4 ml
Bromophenol blue	0.5 ml

Store at 4ºC.

Blocking Buffer

Mix the following:

Fat free skim milk powder	4 g
Tween-20	1 ml
1 X TBS	100 ml

Store at 4 ºC (Short term)

Ponceau S stock solution

Ponceau S	2 g
Trichloroacetic acid	30 g
5-sulfosalicyclic acid	30 g
ddH2O up to	100 ml

For working stock, dilute 1:10. Reusable

10% SDS separating gel (for 2 gels)

Acrylamide mix (37:5:1) (30 %)	5 ml
Tris pH 8.8 (1M)	5.6 ml
SDS (20%)	75 µl
ddH20	4.36 ml
Ammonium persulphate (20 %)	40 µl
TEMED	20 µl

5% SDS stacking gel (for 2 gels)

Acrylamide mix (37:5:1) (30 %)	0.85 ml
Tris pH 6.8 (1M)	0.625 ml
SDS (20%)	25 µl
ddH20	3.5 ml
Ammonium persulphate (20 %)	10 µl
TEMED	2.5 µl

Towbin transfer Buffer

Tris	3.03 g
Glycine	14.42 g
Methanol	200 ml
ddH20 up to	1 L

Store at room temperature.

Solutions for In-house p24 ELISA

Coating solution

5 % bovine serum albumin: Dissolve 5 g in in 1 X TBS

Washing solution

To 1 L of 1 X TBS add 1 ml of Tween-20

PSV inactivation solution

1.25 % Empigen detergent: 1.25 ml of Empigen (35 %) mixed with 33.75 ml of 1 X TBS

Solutions for routine use

0.5 M EDTA

EDTA	93.05 g
NaOH	10 g
ddH20 up to	500 ml

Dissolve the EDTA and NaOH in 400 ml of dH2O, adjust the pH to 8.0 and make up to a final volume of 500 ml.

Phosphate Buffered Saline (PBS)

NaCl	80 g
KCl	2 g
Na2PO4	14.4 g
KH2PO4	2.4 g
ddH20 up to	1 L

Dissolve NaCl, KCl, Na2PO4 and KH2PO4 in 900 ml of ddH_2O. Adjust the pH to 7.4 and make up to 1 L with ddH_2O.